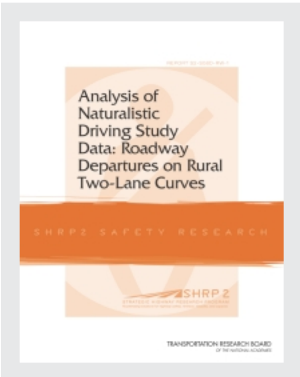


Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves

DETAILS

0 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-43331-0 | DOI 10.17226/22317

AUTHORS

Hallmark, Shauna L.; Oneyear, Nicole; Tyner, Samantha; Wang, Bo; Carney, Cher; and Dan McGehee

The Second

STRATEGIC HIGHWAY RESEARCH PROGRAM

SHRP 2 REPORT S2-S08D-RW-1

Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves

SHAUNA L. HALLMARK, NICOLE ONEYEAR, SAMANTHA TYNER, AND BO WANG
Center for Transportation Research and Education
Institute for Transportation
Iowa State University

CHER CARNEY AND DAN MCGEHEE
University of Iowa

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2015
www.TRB.org

Subject Areas

Data and Information Technology
Highways
Operations and Traffic Management
Safety and Human Factors
Vehicles and Equipment

The Second Strategic Highway Research Program

America's highway system is critical to meeting the mobility and economic needs of local communities, regions, and the nation. Developments in research and technology—such as advanced materials, communications technology, new data collection technologies, and human factors science—offer a new opportunity to improve the safety and reliability of this important national resource. Breakthrough resolution of significant transportation problems, however, requires concentrated resources over a short time frame. Reflecting this need, the second Strategic Highway Research Program (SHRP 2) has an intense, large-scale focus, integrates multiple fields of research and technology, and is fundamentally different from the broad, mission-oriented, discipline-based research programs that have been the mainstay of the highway research industry for half a century.

The need for SHRP 2 was identified in *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life,* published in 2001 and based on a study sponsored by Congress through the Transportation Equity Act for the 21st Century (TEA-21). SHRP 2, modeled after the first Strategic Highway Research Program, is a focused, time-constrained, management-driven program designed to complement existing highway research programs. SHRP 2 focuses on applied research in four areas: Safety, to prevent or reduce the severity of highway crashes by understanding driver behavior; Renewal, to address the aging infrastructure through rapid design and construction methods that cause minimal disruptions and produce lasting facilities; Reliability, to reduce congestion through incident reduction, management, response, and mitigation; and Capacity, to integrate mobility, economic, environmental, and community needs in the planning and designing of new transportation capacity.

SHRP 2 was authorized in August 2005 as part of the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU). The program is managed by the Transportation Research Board (TRB) on behalf of the National Research Council (NRC). SHRP 2 is conducted under a memorandum of understanding among the American Association of State Highway and Transportation Officials (AASHTO), the Federal Highway Administration (FHWA), and the National Academy of Sciences, parent organization of TRB and NRC. The program provides for competitive, merit-based selection of research contractors; independent research project oversight; and dissemination of research results.

SHRP 2 Report S2-S08D-RW-1

ISBN: 978-0-309-27418-0

Copyright Information

Note: SHRP 2 report numbers convey the program, focus area, project number, and publication format. Report numbers ending in "w" are published as web documents only.

Notice

The project that is the subject of this report was a part of the second Strategic Highway Research Program, conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council.

The members of the technical committee selected to monitor this project and review this report were chosen for their special competencies and with regard for appropriate balance. The report was reviewed by the technical committee and accepted for publication according to procedures established and overseen by the Transportation Research Board and approved by the Governing Board of the National Research Council.

The opinions and conclusions expressed or implied in this report are those of the researchers who performed the research and are not necessarily those of the Transportation Research Board, the National Research Council, or the program sponsors.

The Transportation Research Board of the National Academies, the National Research Council, and the sponsors of the second Strategic Highway Research Program do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of the report.

SHRP 2 Reports

Available by subscription and through the TRB online bookstore:

www.mytrb.org/store

Contact the TRB Business Office:
202-334-3213

More information about SHRP 2:

www.TRB.org/SHRP2

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. C. D. (Dan) Mote, Jr., is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Victor J. Dzau is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. C. D. (Dan) Mote, Jr., are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

ACKNOWLEDGMENTS

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials. It was conducted in the second Strategic Highway Research Program (SHRP 2), which is administered by the Transportation Research Board of the National Academies. The project was managed by Kenneth L. Campbell, Chief Program Officer for SHRP 2 Safety, and James Hedlund, SHRP 2 Special Consultant for Safety Coordination.

The authors thank the Virginia Tech Transportation Institute (VTTI) for providing the Naturalistic Driving Study data and access to the secure data enclave, as well as for providing expertise and assistance in working with the data. The authors also thank the Center for Transportation Research and Education (CTRE) at Iowa State University for providing the Roadway Information Database data.

FOREWORD

James Hedlund, *SHRP 2 Special Consultant, Safety Coordination*

The SHRP 2 Naturalistic Driving Study (NDS) was the largest and most comprehensive study of its kind ever undertaken. Its central goal was to produce unparalleled data from which to study the role of driver performance and behavior in traffic safety and how driver behavior affects the risk of crashes. Such research involves understanding how a driver interacts with and adapts to the vehicle, the traffic environment, roadway characteristics, traffic control devices, and other environmental features. After-the-fact crash investigations can only provide this information indirectly. The NDS data recorded how drivers really drove and what they were doing just before they crashed or almost crashed.

The Roadway Information Database (RID), created in parallel with the NDS, contains detailed roadway data collected on more than 12,500 centerline miles of highways in and around the six study sites, about 200,000 highway miles of data from the highway inventories of the six study states, and additional data on crash histories, traffic and weather conditions, work zones, and ongoing safety campaigns in the study sites.

The NDS and RID data can be linked to associate driving behavior with the roadway environment. The data will be used for years to come for developing and evaluating safety countermeasures designed to prevent or reduce the severity of traffic crashes and injuries.

The NDS collected data from more than 3,000 male and female volunteer passenger-vehicle drivers, aged 16 to 98, during a 3-year period. Most drivers participated from 1 to 2 years. It was conducted at one site in each of six states: Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington. Data collected included vehicle speed, acceleration, and braking; vehicle controls, when available; lane position; forward radar; and video views forward, to the rear, and on the driver's face and hands. The NDS data file contains about 50 million vehicle miles, 5 million trips, more than 3,900 vehicle years, and more than 1 million hours of video—a total of about 2 petabytes of data.

Four contracts were awarded in 2012 under SHRP 2 Safety Project S08, Analysis of the SHRP 2 Naturalistic Driving Study Data, to study specific research questions using the early NDS and RID data. An open competition solicited proposals to address topics of the contractor's own choosing that would have direct safety applications and that would

- Lead to real-world applications and safety benefits (theoretical knowledge without potential applications was not a priority);
- Be broadly applicable to a substantial number of drivers, roadways, or vehicles in the United States; and
- Demonstrate the use of the unique NDS data (i.e., similar results could not be obtained from existing nonnaturalistic data sets).

In addition to these goals, SHRP 2 expected the projects to serve as both pilot testers and advisers. As they conducted these first substantial NDS and RID analyses, these studies' experienced researchers would discover valuable insights on a host of both pitfalls and opportunities that others should know about when they use the data.

The four projects began in February 2012 and were conducted in two phases. In Phase 1, which concluded in December 2012, contractors obtained an initial set of data, tested and refined their research plans, and developed detailed plans for their full analyses. Three projects successfully completed this proof of concept and were selected for Phase 2. These three projects obtained and analyzed a much richer, though still preliminary, data set and reported their results in July 2014. This report, *Analysis of Naturalistic Driving Study Data: Roadway Departures on Rural Two-Lane Curves,* documents one of the three projects.

These projects were conducted while the NDS and RID data files were being built. This circumstance imposed constraints that substantially affected the researchers' work. The constraints included the following:

- *Sample size.* In summer 2013, when the projects requested full data sets, the NDS data file was only 20% to 30% complete. As a result, each project could only obtain a fraction of the trips of interest now available in the full NDS data.
- *RID not complete and not linked to the NDS.* Projects based on roads of specific types or locations could not identify these roads from the RID but instead had to use Google Earth or a similar database to identify them. They then obtained trips of interest by using searches through the NDS that were less efficient than will be possible when the NDS and RID are linked.
- *Data processing.* Some data, such as radar, had not been processed from their raw state to a form where they were fully ready for analysis.
- *Data quality.* NDS data are field data, and field data are inherently somewhat messy. At the time these projects obtained their data, some data had not been quality controlled, and some characteristics of the data were not yet well understood.
- *Tools for data users.* Not all crashes and near crashes had been identified, and a separate small data set containing only crashes, near crashes, and baseline exposure segments had not been built. In addition, a small trip summary file containing key features of each trip had not been built. Users can conduct initial analyses on many subjects quickly and easily using a trip summary file.
- *Other demands on data file managers.* The first priority for the NDS manager, Virginia Tech Transportation Institute (VTTI), and the RID manager, Iowa State University's Center for Transportation Research and Education (CTRE), was to complete data processing and quality control. Field data were being ingested continually. Data delivery for users was sometimes delayed because of these demands on their resources.

These issues are being resolved in 2014. The NDS and RID data are complete and are being linked. Data processing and quality control are being completed. Crash and near-crash files and trip summary files are being built.

If this project and the other two were to begin in 2015, each would have more data and would obtain the data far more easily and quickly. Readers should keep these constraints in mind as they read this report. Despite working under these constraints, the three NDS projects have produced valuable new insights into important traffic safety issues that will help reduce traffic crashes and injuries.

For an overview of the study, see the following article: K. L. Campbell, The SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety, *TR News,* No. 282, September–October 2012, pp. 30–35. Additional details may be found at the study's InSight website: https://insight.shrp2nds.us/.

CONTENTS

1 Executive Summary
1 Project Objectives
1 Research Questions Addressed
1 Data Collection and Reduction
2 Crash Surrogates
2 Results and Discussion

6 CHAPTER 1 SHRP 2 Naturalistic Driving Study Background
6 The First SHRP 2 NDS Analysis Projects
7 Constraints of the First SHRP 2 NDS Studies

8 CHAPTER 2 Introduction
8 Background
8 Rationale for Research
9 Objectives
9 Research Questions Addressed

10 CHAPTER 3 Data Collection
10 Identification of Data Needs
10 Data Requests
15 Summary of Data Received and Limitations

17 CHAPTER 4 Data Reduction
17 Reduction of Roadway Variables
17 Reduction of Vehicle, Traffic, and Environmental Variables
20 Reduction of Kinematic Driver Characteristics
22 Summary of Data Limitations

26 CHAPTER 5 Selection of Crash Surrogates
26 Identification of Possible Roadway Departure Crash Surrogates
27 Description of Selected Crash Surrogates

29 CHAPTER 6 Analysis for Research Question 1
29 Data Sampling and Variables Used for Research Question 1
30 Methodology for Defining Curve Area of Influence
31 Results for Research Question 1
35 Summary and Discussion

37 CHAPTER 7 Analysis for Research Question 2
37 Data Sampling and Segmentation Approach for Research Question 2
38 Variables Used for Research Question 2
38 Description of Analytical Approach for Research Question 2
40 Results for Research Question 2
42 Summary and Implications

44 CHAPTER 8 Analysis for Research Question 3
44 Data Sampling and Modeling Approach for Research Question 3
45 Variables Used for Research Question 3
46 Description of Analytical Approach and Results for Research Question 3
49 Summary of Crash/Near-Crash Events
49 Summary and Discussion

51 CHAPTER 9 Analysis for Research Question 4
51 Description of Analytical Approach for Research Question 4
51 Data Sampling and Segmentation Approach for Research Question 4
52 Variables Used for Research Question 4
52 Results for Research Question 4
54 Summary and Implications

56 CHAPTER 10 Summary and Recommendations
56 Summary
56 Discussion and Recommendations for Countermeasures

59 References

61 Appendix A. Methodology for Reducing Roadway Data

69 Appendix B. Data Reduction Method for Coding Driver Glance Location and Distraction

72 Appendix C. Curve Area of Influence Model Results

Executive Summary

Project Objectives

Rural curves are known to pose a significant safety problem, but the interaction between the driver and roadway environment is not well understood. Thus, the objective of this research was to assess the relationship between driver behavior and characteristics, roadway factors, environmental factors, and the likelihood of roadway departures on rural two-lane curves.

To accomplish this, data from the second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) and Roadway Information Database (RID) were used to develop relationships between driver, roadway, and environmental characteristics and the risk of a roadway departure on curves.

The project focused on rural two-lane curves on paved roadways. Only paved roadways were included because the machine vision application used in the lane-tracking system does not function well when lane lines or obvious discontinuities between the lane and shoulder surface are not present. Rural was defined as one or more miles outside an urban area. Additionally, only roadways posted at 64 km/h to 97 km/h (40 mph to 60 mph) were included.

Research Questions Addressed

This research was tailored to address four fundamental research questions:

1. What defines the curve area of influence?
2. What defines normal behavior on curves?
3. What is the relationship between driver distractions; other driver, roadway, and environmental characteristics; and risk of roadway departure?
4. Can lane position at a particular state be predicted as a function of position in a prior state?

Each question addresses the problem from a different perspective. As a result, a different methodology was proposed for each, as described in the corresponding sections.

Data Collection and Reduction

Chapter 3 summarizes how Institutional Review Board (IRB) approval and data requests were completed; it also describes data reduction. The team manually identified rural curves in Florida, New York, Indiana, Pennsylvania, and North Carolina based on information about where trips were likely to have occurred. Segments were provided to Virginia Tech Transportation Institute (VTTI) staff, who identified trips through those segments.

Roadway, environmental, and operational characteristics were extracted as described in Chapter 4. Site visits were made to the VTTI secure data enclave to reduce driver glance location and distraction for each trace.

Crash Surrogates

The use of crash surrogates was necessary because only one crash and three near crashes were available at the time this research was conducted. Chapter 5 discusses the rationale for selection of the identified crash surrogates. A number of potential crash surrogates were considered against the data available and the expected accuracy of relevant variables in the NDS data (e.g., lane position, forward radar, vehicle position). Lane offset was the best crash surrogate, but lane offset was not reliable in a number of traces. As a result, it was used for Research Questions 2 and 4, resulting in a smaller sample of data for those research questions.

Because offset was not reliable in a number of traces, the research team determined that encroachments would be the best crash surrogate for Research Question 3. A *right-side encroachment* is defined as the right side of the vehicle crossing the right lane line, and a *left-side encroachment* is defined as the left side of the vehicle crossing the centerline.

Results and Discussion

Since four fundamental research questions were addressed, a different methodology was developed specific to each, as outlined in Chapters 6 through 9. In addition to the analytical method, these chapters discuss the data sampling and segmentation approach, general variables considered, results, and implications. The following sections provide a brief summary of findings and implications for each research question.

Research Question 1

Answering Research Question 1 entailed understanding at what point drivers begin reacting to the presence of a curve upstream of the curve. Understanding where drivers begin to react to the curve is important for placement of traffic control and countermeasures. A better understanding can also help agencies determine optimal placement of advance signing and other countermeasures. Research Question 1 was also used to indicate the curve area of influence for Research Questions 2, 3, and 4.

Time series data were modeled using regression and Bayesian analysis. Results indicate that, depending on radius of curve, drivers begin reacting to the curve 164 m to 180 m (538.1 ft to 590.6 ft) upstream of the point of curvature. These results were compared with sign placement guidelines in the *Manual on Uniform Traffic Control Devices* (Federal Highway Administration 2009), and it was determined that the guidelines are appropriately set based on where drivers actually react to the curve.

Research Question 1 also found that drivers begin reacting to the curve sooner for curves with larger radii than for curves with smaller radii. Drivers may not be able to gauge the sharpness of the curve, or sight distance issues may be a concern for sharper curves.

This suggests that use of countermeasures—such as chevrons or raised pavement markings (RPMs)—that better delineate the curve may provide better advance information for drivers. It should be noted that the model only identified the point at which drivers reacted to the curve. This research question did not attempt to answer whether the reaction point was sufficient for drivers to successfully negotiate the curve.

Research Question 2

Research Question 2 developed conceptual models of curve driving to assess changes in metrics as the driver negotiates the curve. Understanding how a driver normally negotiates a curve provides insight not only into how characteristics of the roadway, driver, and environment potentially influence driving behavior, but also into areas that can lead to roadway departures. Knowing how much drivers normally deviate in their lane, as well as how they choose their speed, could potentially have implications on policy or design.

Data for several positions upstream and along the curve were sampled from the time series data. Models were developed for lane position and speed for both inside (right-hand curve from the perspective of the driver) and outside (left-hand curve from the perspective of the driver). Lane position was modeled as the offset of the center of the vehicle from the center of the lane. Models were developed using generalized least squares.

Results indicate that lane position within the curve is influenced by lane position upstream of the curve. The models developed for offset of lane centerline found that drivers who were distracted or who glanced away from the roadway tended to shift away from the center of the lane. When driving on the inside of the lane, a driver who was distracted at a particular point within the curve tended to shift 0.14 m to the right by the next point in the curve. When driving on the outside (left-hand curve), a driver who engaged in a non-roadway-related glance at a particular location within the curve was expected to move to the left, or toward the centerline by 0.13 m at that same point. This confirms the role of distraction in lane keeping.

Additionally, the models found that drivers on the inside of a curve tended to move more to the right at the center of curve, while drivers on the outside of curves were at the furthest point from the lane centerline at the beginning of the curve. As a result, drivers may be particularly vulnerable to roadway departures at certain points in the curve negotiation process. These results suggest that countermeasures such as rumble strips, paved shoulders, and high-friction treatments may ameliorate the consequences of variations in lane position through the curve.

Additionally, the lane offset models indicate that age and nighttime driving are factors in driver lane position.

The model for speeding in the curve found that if drivers are speeding in the upstream, they will also speed in the curve. Drivers of sport utility vehicles (SUVs) and pick-up trucks traveled on average 2.1 km/h (1.3 mph) faster than drivers of passenger vehicles.

Speeds were predicted to be 0.9 km/h (0.5 mph) lower for each additional 10 years in age for a driver, and drivers engaged in a non-roadway-related glance are expected to travel 5.3 km/h (3.3 mph) slower than drivers who do not engage in a non-roadway-related glance. This suggests that drivers whose attention is focused away from the roadway do not maintain longitudinal control.

The results indicate that distractions/nonroadway glances affect lateral and longitudinal control. Although drivers are more likely to travel at slower speeds, they are more likely to vary within their lane.

The models also confirm that speed plays an important role in curve negotiation. This suggests that effective speed management countermeasures will have an impact on curve negotiation. Speed management countermeasures include better delineation of the curve (i.e., chevrons, edge lines, post-mounted delineators) so that drivers can better estimate the sharpness of the curve. Other measures, such as transverse speed markings or speed feedback signs, target drivers who are traveling over the speed limit.

Research Question 3

Research Question 3 addressed how driver behavior in conjunction with roadway and environmental factors affect the likelihood of a roadway departure on rural two-lane curves. Four different models were developed using multivariate logistic regression. Two models evaluated the probability (odds) of a right-side or left-side encroachment based on driver, roadway, and environmental characteristics. Two additional models evaluated the probability that a driver would exceed the advisory speed if present or posted speed limit if not present at the curve entry by 8 km/h and 16 km/h or more (5 mph and 10 mph or more). Data were aggregated to the event level.

The model for right-side encroachments indicates that the probability increases as drivers spend less time glancing at the forward roadway. The results also indicate that a right-side lane departure is 6.8 times more likely on the inside of a curve compared with the outside of the curve. Lane departures are slightly more likely (1.3 times) for curves with any type of curve advisory sign (including W1-6). It is unlikely that the presence of a warning sign itself increases the probability. Rather, it is likely that advisory signs are more likely to be present on curves of a certain type (i.e., those with sight distance issues, sharper curves), and encroachments are also more likely for those road types. Additionally, the results suggest that the simple presence of curve warning signs may not mitigate roadway departures.

A statistically significant but small correlation exists between radius of curve and probability of a right-side encroachment. Drivers were 0.33 times less likely to have a right-side encroachment on roadways with a guardrail. Presence of a guardrail decreased the probability of a right-side encroachment. The purpose of a guardrail is to mitigate the consequences of a driver leaving the roadway rather than to keep the driver from leaving the roadway. Consequently, a guardrail in and of itself does not mitigate roadway departures. The presence of a guardrail may suggest to the driver that roadway conditions are less safe, resulting in better driver attention. Additionally, few delineation countermeasures (e.g., chevrons) were present in the curves included in the analysis. As a result, a guardrail may provide some delineation of the curve, which provides feedback to the driver about the sharpness of the curve.

The model for left-side encroachments indicates that males are more than four times more likely to have a left-side lane departure, and drivers traveling on the inside of the curve are 0.1 times less likely to have a left-side encroachment than drivers traveling on the outside of the curve. The impact of radius was statistically significant but minor.

The probability that a driver will be 8 km/h (5 mph or more) over the posted/advisory speed is higher when the driver is younger and has a higher average speed upstream and when edge line markings are obscured or not present. The amount of time a driver spends following another vehicle, presence of lower visibility conditions, and presence of paved shoulders and RPMs decrease the probability that he or she will enter the curve 5 mph or more over the posted/advisory speed.

The probability that a driver will be 16 km/h (10 mph or more) over the posted/advisory speed is higher when the driver has a higher average speed upstream. The probability is lower when the average glance at roadway-related tasks is longer and when paved shoulders and RPMs are present.

Results from the right-side encroachment and speed models suggest that better curve delineation may allow drivers to better gauge upcoming changes in roadway geometry, resulting in better speed selection and decreased risk of a roadway departure, and may help decrease speed. Delineation countermeasures include chevrons, the addition of reflective panels to existing chevron posts, reflective barrier delineation, RPMs, post-mounted delineators, edge lines, and wider edge lines.

The speed models suggest that driver age and upstream speed have a significant impact on drivers' speed within a curve. As a result, speed management countermeasures that affect tangent speed will also decrease curve speeds. The results also indicate that speed management is appropriate to get drivers' attention before entering a curve. Countermeasures specifically targeted to reduce speed on curves include dynamic speed feedback signs, on-pavement curve warning signs, and flashing beacons.

Research Question 4

Research Question 4 focused more specifically on driver response to changing roadway characteristics and traffic conditions. Time series models were developed to incorporate the dynamic process of information acquisition and response as a driver negotiates a curve. The analysis evaluated the influence of roadway geometries or traffic conditions on drivers' lane-keeping behavior. For example, drivers on a rural two-lane roadway tend to have larger lane deviation from the centerline when there is an oncoming vehicle.

Two types of dynamic linear models (DLMs) were built in this study to describe and explain the curve negotiation process: DLM with intervention analysis and DLM with autoregression and moving average (ARMA). The DLM with intervention analysis was mainly used for explanatory purposes, relating lane offset to curve characteristics and traffic conditions. The DLM with ARMA was mainly used for forecasting purposes, which could be used for a roadway departure warning system.

Only limited data were used in the analyses, given that the objective was to demonstrate the utility of the approach. Results indicate that lane position can successfully be modeled as a function of vehicle position in a prior state and as a function of other characteristics such as position within the curve or presence of oncoming vehicles. The methodology shows promise for use in development of roadway departure crash warning systems.

CHAPTER 1

SHRP 2 Naturalistic Driving Study Background

The second Strategic Highway Research Program (SHRP 2) conducted the largest and most comprehensive naturalistic driving study (NDS) ever undertaken. The study collected data from more than 3,000 male and female volunteer passenger-vehicle drivers, aged 16 to 98, during a 3-year period, with most drivers participating from 1 year to 2 years. The study was conducted in six sites, one each in Florida, Indiana, New York, North Carolina, Pennsylvania, and Washington. Data collected include vehicle speed, acceleration, and braking; vehicle controls when available; lane position; forward radar; and video views forward, to the rear, and on the driver's face and hands. The NDS data file contains about 50 million vehicle miles, 5 million trips, more than 3,900 vehicle years, and more than 1 million hours of video—a total of about 2 petabytes of data.

In parallel, the Roadway Information Database (RID) contains detailed roadway data collected on more than 12,500 centerline miles of highways in and around the study sites, about 200,000 highway miles of data from the highway inventories of the six study states, and additional data on crash histories, traffic and weather conditions, work zones, and ongoing safety campaigns in the study sites. The NDS and RID data can be linked to associate driving behavior with the roadway environment.

Campbell (2012) provides an excellent overview of the study. Additional details may be found at the study's InSight website (https://insight.shrp2nds.us/).

The study's central goal is to produce unparalleled data from which to study the role of driver performance and behavior in traffic safety and how driver behavior affects the risk of crashes. This involves understanding how the driver interacts with and adapts to the vehicle, the traffic environment, roadway characteristics, traffic control devices, and other environmental features. After-the-fact crash investigations can do this only indirectly. The NDS data record how drivers really drive and what they are doing just before they crash or almost crash. The NDS and RID data will be used for years to come to develop and evaluate safety countermeasures designed to prevent or reduce the severity of traffic crashes and injuries.

The First SHRP 2 NDS Analysis Projects

Four contracts were awarded in 2012 under SHRP 2 Project S08, Analysis of the SHRP 2 Naturalistic Driving Study Data, to study specific research questions using the early SHRP 2 NDS and RID data. An open competition solicited proposals to address topics of the contractor's own choosing that would have direct safety applications. The request for proposals required proposals that would

- Lead to real-world applications and safety benefits (theoretical knowledge without potential applications was not a priority);
- Be broadly applicable to a substantial number of drivers, roadways, and/or vehicles in the United States; and
- Demonstrate the use of the unique NDS data (i.e., similar results could not be obtained from existing nonnaturalistic data sets).

In addition to these goals, SHRP 2 expected these projects to serve as both pilot testers and advisers. As they conducted these first substantial NDS and RID analyses, these studies' experienced researchers would discover valuable insights on a host of both pitfalls and opportunities that others should know about when they use the data. That experience and advice can be found on the study's InSight website (https://insight.shrp2nds.us/).

The four projects began in February 2012 and were conducted in two phases. In Phase 1, which concluded in December 2012, the four contractors each obtained an initial set of data, tested and refined their research plan, and developed a detailed plan for their full analyses. Three projects, of which this study is one, successfully completed this proof-of-concept phase and were selected for the full Phase 2. These three projects obtained and analyzed a much richer, though still preliminary, data set and reported their results in July 2014.

Constraints of the First SHRP 2 NDS Studies

These projects were conducted while the NDS and RID data files were being built. This circumstance imposed constraints that substantially affected the researchers' work. The constraints included the following:

- *Sample size.* In summer 2013, when the projects requested their full data sets, the NDS data file was only 20% to 30% complete. As a result, each project could obtain only a fraction of the trips of interest now available in the full NDS data.
- *RID not complete and not linked to the NDS.* Projects based on roads of specific types or locations could not identify those roads from the RID but instead had to use Google Earth or some similar database to identify them. Researchers then obtained trips of interest using less efficient searches through the NDS than will be possible when the NDS and RID are linked.
- *Data processing.* Some data, such as radar, had not been processed from their raw state to a form where they were fully ready for analysis.
- *Data quality.* NDS data are field data, and field data are inherently somewhat messy. When these projects obtained their data, some data had not been quality controlled and some characteristics of the data were not yet well understood.
- *Tools for data users.* Not all crashes and near crashes had been identified, and a separate small data set containing only crashes, near crashes, and baseline exposure segments had not been built. In addition, a small trip summary file containing key features of each trip had not been built. Users can conduct initial analyses on many subjects quickly and easily using a trip summary file.
- *Other demands on data file managers.* The first priority for the NDS manager, Virginia Tech Transportation Institute (VTTI), and the RID manager, Iowa State University's Center for Transportation Research and Education (CTRE), was to complete data processing and quality control. Field data were being ingested continually. Data delivery for users sometimes was delayed due to these demands on their resources.

These issues are being resolved in 2014. The NDS and RID data are complete and are being linked. Data processing and quality control are being completed. Crash and near-crash files and trip summary files are being built. If this project and the other two were to begin in 2015, each would have more data and would obtain the data far more easily and quickly. Readers should keep these constraints in mind as they read this report. Despite working under these constraints, this project and the other two NDS projects have produced valuable new insights on important traffic safety issues that will help reduce traffic crashes and injuries.

CHAPTER 2

Introduction

Background

The Federal Highway Administration (2009) estimates that 58% of roadway fatalities are roadway departures, while 40% of fatalities are single-vehicle run-off-road (SVROR) crashes. Addressing roadway departure crashes is therefore a priority for national, state, and local roadway agencies; horizontal curves are of particular interest because they have been correlated with overall increased crash occurrence. Glennon et al. (1985) reported that curves have approximately three times the crash rate of tangent sections, and Preston (2009) reported that 25% to 50% of severe road departure crashes in Minnesota occurred on curves, even though they account for only 10% of the system mileage. Farmer and Lund (2002) found that the odds of having a rollover on a curved section were 1.42 to 2.15 times greater than the odds of having a rollover on a straight section. The majority of crashes on curves involve roadway departures. A total of 76% of curve-related fatal crashes are single vehicles leaving the roadway and striking a fixed object or overturning. Another 11% of curve-related crashes are head-on collisions (AASHTO 2008).

Curve-related crashes have a number of causes, including roadway and driver factors. Degree of curve or radius of curve is the roadway factor most cited in the literature as having an impact on crash risk (Luediger et al. 1988; Miaou and Lum 1993). Other factors that have been correlated to the frequency and severity of curve-related crashes include length of curve, type of curve transition, lane and shoulder widths (Zegeer et al. 1991), preceding tangent length (Milton and Mannering 1998), presence of spirals (Council 1998), grade (Fink and Krammes 1995), and required speed reduction between the tangent and curve.

Driver error on horizontal curves is often due to inappropriate speed selection, which results in an inability to maintain lane position. FHWA estimates that approximately 56% of run-off-road (ROR) fatal crashes on curves are speed related. Distracting tasks such as radio tuning or cell phone conversations can draw a driver's attention away from speed monitoring, changes in roadway direction, lane keeping, and detection of potential hazards (Charlton 2007). Other factors include sight distance issues, fatigue, or complexity of the driving situation (Charlton and DePont 2007; Charlton 2007). McLaughlin et al. (2009) evaluated ROR events in VTTI's 100-car naturalistic driving study and found that distraction was the most frequently identified contributing factor, along with fatigue, impairment, and maneuvering errors.

Environmental factors, such as the roadway surface condition, and vehicle factors, such as the center of gravity, also have an impact on a driver's ability to safely negotiate a curve. McLaughlin et al. (2009) found that ROR events were 1.8 times more likely on wet roads than dry, 7.0 times more likely on roads with snow or ice than dry roads, and 2.5 times more likely in nighttime than daytime conditions.

Studies of roadway factors—such as degree of curve, presence of spirals, or shoulder width and type—have provided some information regarding which curve characteristics are the most relevant, but information is still lacking. In addition, little information is available that identifies driver behaviors that contribute to curve crashes. As a result, a better understanding of how drivers interact with various roadway features and countermeasures will provide valuable information to highway agencies in determining how resources can best be allocated to maximize driver performance and reduce crashes.

Rationale for Research

Although some studies have assessed the relationship between crash risk and driver and roadway characteristics, the contributory factors have not been well established. This is primarily because crash data typically have only a limited number of roadway variables, and driver behavior leading up to a crash can only be inferred by the reporting police officer.

This lack of understanding makes it difficult to design, select, and apply the appropriate countermeasures, given that safety professionals do not really understand how drivers are failing to interact with the roadway. Most studies that have evaluated countermeasures are based on crash or speed reductions. The results of these studies only demonstrate that a countermeasure works or does not work. They do not explain why the countermeasure works. For instance, several studies have examined the use of wider pavement markings or raised pavement markings (Donnell et al. 2006) to reduce crashes in curves. Although the treatments have shown some promise, why they are effective is not well understood. The treatment may be successful for several reasons: (1) the treatment provides better delineation so that a driver is better able to judge the sharpness of the curve; (2) the treatment may simply get the driver's attention; or (3) the treatment may cause the lane to appear narrower, causing the driver to feel more constrained and resulting in lower speeds. Because it is unclear why the treatment is effective, it is unclear whether applying a particular countermeasure at a different curve will also be effective, and the understanding of the problem is insufficient to support the design and selection of alternative treatments.

A better understanding of the interaction between driver characteristics and curve negotiation needs can lead to better design and application of countermeasures. For instance, if older drivers have the hardest time with curve negotiation because they are less likely to see visual cues, the best solution might be larger chevrons. Conversely, a solution geared toward younger drivers might include more closely spaced chevrons to help drivers gauge the sharpness of the curve. Distracted drivers might require another solution, such as a tactile cue from transverse rumble strips.

Objectives

Rural curves pose a significant safety problem, and the interaction between the driver and roadway environment is not well understood. To address this knowledge gap, this research aimed to assess the relationship between driver characteristics and behavior, roadway factors, environmental factors, and the likelihood of roadway departures on rural two-lane paved curves.

To accomplish this objective, the second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study (NDS) and Roadway Information Database (RID) were used to develop models that explore how drivers interact with the roadway environment and that identify the conditions present both when a driver does not successfully negotiate a rural curve and when successful negotiation occurs. These conditions include driver, roadway, and, to limited extent, environmental conditions.

Most highway agencies are proactive in implementing a range of countermeasures to reduce roadway departures on curves and in other areas. However, agencies have only limited information about the effectiveness of different countermeasures. The results of this research will provide more information about the specific roadway features that are correlated to increased risk of roadway departure. The results will also provide valuable information about how drivers interact with roadway features and the impact that that interaction has on the effectiveness of countermeasures. This information will allow agencies to make better decisions about countermeasure selection.

This research focused on two-lane rural curves. Addressing roadway departures on all curves is important; however, the SHRP 2 NDS encompassed almost 5 million trips, and—given time and resources constraints—researchers could not consider all curve-related roadway departure scenarios. Rural two-lane roads were prioritized and selected because of the disproportionate number of roadway departure crashes experienced on these roads (Garder 2006; Fitzpatrick et al. 2002).

Only paved roadways were included because the machine vision application does not function well when lane lines or obvious discontinuities between the lane and shoulder surface are not present. *Rural* was defined as one or more miles outside an urban area. Additionally, only roadways posted at 64 km/h to 97 km/h (40 mph to 60 mph) were included.

Research Questions Addressed

This main research question addressed was the following: What is the relationship between driver distraction, other driver characteristics, roadway characteristics, environmental characteristics, and risk of roadway departure?

The data that were obtained and reduced allowed researchers to explore driver behavior on curves in several additional ways. Each way offers a different perspective, and each was posed as an individual research question. As a result, a different methodology was proposed for each, as described in the corresponding sections.

This research was tailored to address four fundamental research questions:

1. What defines the curve area of influence?
2. What defines normal behavior on curves?
3. What is the relationship between driver distraction; other driver, roadway, and environmental characteristics; and risk of roadway departure?
4. Can lane position at a particular state be predicted as a function of position in a prior state?

As described in Chapter 4, *distraction* for the purposes of this research was defined as engagement in a non-driving-related activity while the driver was glancing in a location other than the forward roadway, or "eyes-off-roadway."

CHAPTER 3

Data Collection

This chapter describes the process for obtaining and reducing the different NDS data sets used in the present analysis. A further description of the data reduction effort for each specific analysis is included within each corresponding section.

Identification of Data Needs

Before requesting data, the research team identified the desired driver, roadway, and vehicle characteristics necessary to answer the stated research questions. To accomplish this, the team conducted a number of literature reviews regarding factors related to rural roadway departures and researched the impact of various countermeasures. This information and the team's expertise were used to develop a list of roadway, driver, and environmental data elements necessary to answer the stated research questions. An in-depth summary can be found in Hallmark et al. (2011). In addition, the team completed an assessment of the SHRP 2 NDS and RID data to determine the likely sources of data and the accuracy of that data. This list of desired data elements was used to guide the data requests, which are described in the following section.

Data Requests

New Institutional Review Board (IRB) and data-sharing agreements were necessary to initiate data requests. Both the Center for Transportation Research and Education at Iowa State University (CTRE/ISU) and University of Iowa (UI) teams obtained the appropriate IRB approval from their respective home institutions and then submitted data-sharing agreements to the Virginia Tech Transportation Institute (VTTI).

Both the NDS and RID data were still being collected at the time this analysis was being conducted. In addition, the NDS and RID data had not yet been linked at the time the data requests were initiated. As a result, it was necessary for the CTRE/ISU team to manually identify potential curves of interest using the method described below.

The team focused on Florida (FL) for Phase 1 because NDS and roadway data collection were the most advanced in that area at the time Phase 1 commenced. Phase 2 included North Carolina (NC), Indiana (IN), New York (NY), and Pennsylvania (PA) because those states had the greatest share of rural roadway data. Table 3.1 provides a summary of the data that were ultimately collected for the RID by state.

To identify potential curves of interest, the project team made use of weighted section maps that VTTI prepared during the early stages of the NDS data collection to help the team working on SHRP 2 Safety Project S04A, Roadway Information Database Development, to focus mobile mapping on roadway sections where NDS trips were occurring.

A data-sharing agreement and data request were made so that the weighted section maps could be used to identify curves of interest. The team manually reviewed the weighted section maps and the RID and identified segments of rural two-lane paved roadways with curves. *Rural* was defined as approximately 1 mile from an urban or built-up area. A *segment* consisted of a continuous stretch of roadway with no major changes in roadway cross section. Each segment included one or more curves. Each segment also included a tangent distance of at least 0.5 miles upstream or downstream of the first or last curve included in the respective segment.

A buffer was created around each segment so that trips through the segment could be identified by VTTI. A buffer was used so that vehicle activity passing along the corresponding roadway section could be "clipped" out using GIS overlay functions. Trips through a buffer are referred to in this report as *traces* because the term *trip* is used in the NDS to indicate a journey from origin to destination. Traces are thus portions of a trip. An example of a buffer section and corresponding traces is shown in Figure 3.1.

Identified segments were evaluated and were removed from further consideration when they included turning or passing lanes in the curve, stop or signal controlled intersections, or a

Table 3.1. Rural/Urban Split for RID Data Collection by State

Study State	Miles Collected for RID	Rural/Urban Split
FL	4,366	45% rural/55% urban
IN	4,635	64% rural/36% urban
NC	4,558	59% rural/41% urban
NY	3,570	68% rural/32% urban
PA	3,670	83% rural/17% urban
WA	4,277	31% rural/69% urban

Table 3.2. Location of Buffer Segments

Study State	Buffers	Curves
IN	80	375
NY	71	173
NC	20	58
PA	32	101
FL	14	32
Total	**217**	**739**

significant portion in which the surrounding area appeared to be more urban than rural. Fourteen buffers and 32 curves were identified in Florida in Phase 1, and the remaining were identified in Phase 2. A total of 217 buffers and 739 curves were identified, as shown in Table 3.2.

Segment locations identified for Phases 1 and 2 (Florida, New York, Pennsylvania, North Carolina, and Indiana) are shown in Figure 3.2.

A sampling plan was developed after initial analyses in Phase 1, and a total request for about 1,000 traces was planned for Phase 2. This request was based on an estimate of an ideal sample size for the statistical methods identified for the research questions and was balanced against what the CTRE/ISU and UI teams could reasonably accomplish in Phase 2. The resources needed for VTTI to complete the request were also considered in selection of the sampling size.

Data were requested in three phases, as described in the following sections. Each trace represents one trip through one buffer segment by one driver. Only data for which the driver traversed the entire segment were requested because the driver may have turned onto or off of the segment.

Source: World_Imagery (Esri, DigitalGlobe, GeoEye, i-cubed, USDA, USGS AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community).

Figure 3.1. Example of traces through buffer segment.

VTTI typically first provided a data file with output from the in-vehicle data acquisition system (DAS) along with static driver/vehicle characteristics such as age, gender, and vehicle type. The CTRE/ISU team reviewed the raw DAS file to determine whether sufficient data were available for variables of interest; VTTI then provided a video with a forward and rear view of the roadway only for these traces. The data elements to be provided were specified in the data request.

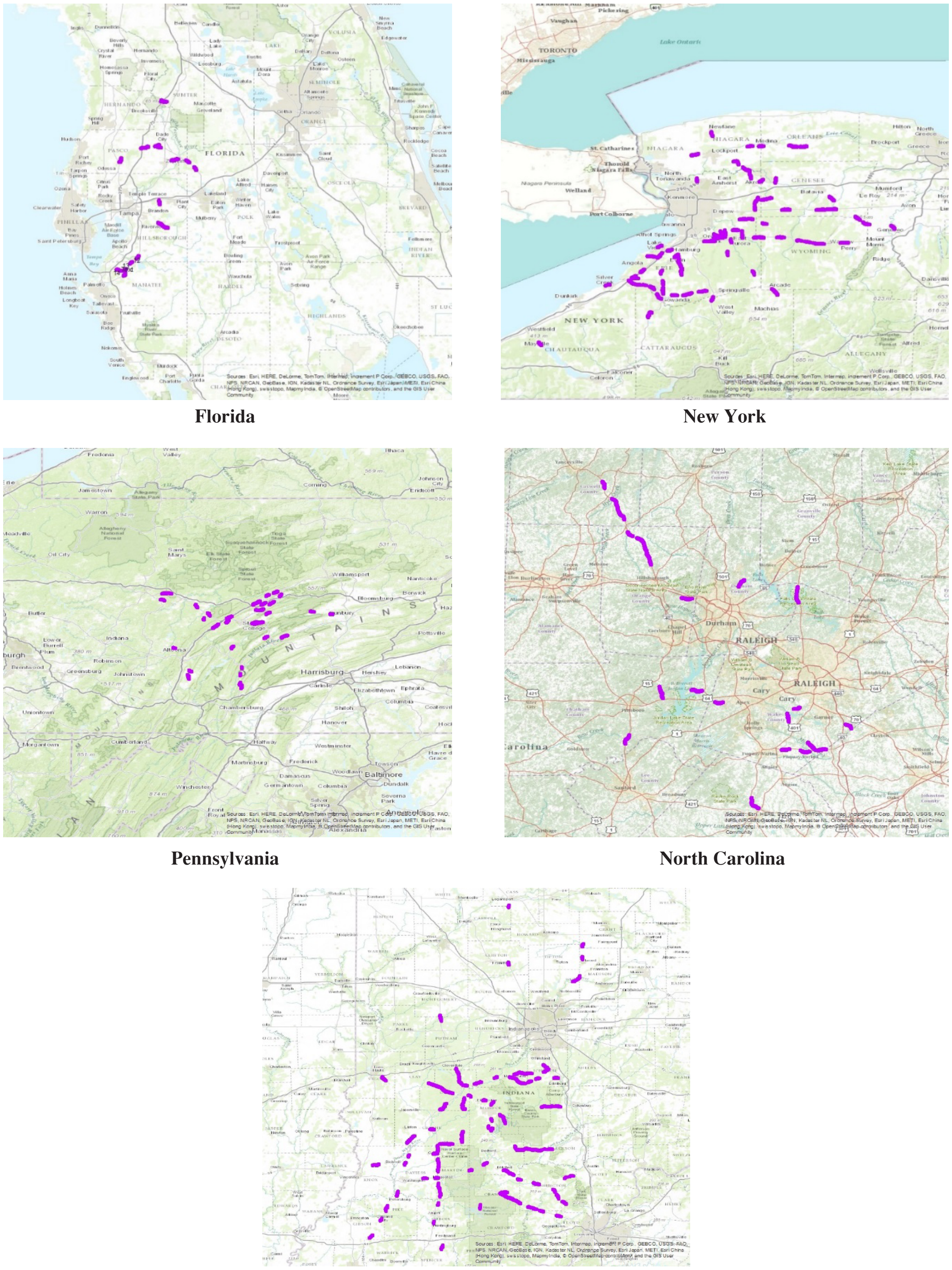
Driver face and steering wheel position/hand position videos were viewed and reduced by the team at the VTTI secure data enclave, as described in Chapter 4.

Information such as vehicle speed, acceleration, pedal position, and wiper blade position were provided in the data file. Additional information about available variables is provided in Chapter 4, but an example of a raw data file is shown in Table 3.3. Data were provided at 10 Hz (0.1 s intervals). GPS location was also provided so that the data could be imported into a geographic information system (GIS) program and overlaid with the RID and aerial imagery. These data are referred to as time series DAS data. Video and time series data are linked using timestamps.

An example of the video views available is shown in Figure 3.3. The forward and rear roadway video views were provided along with the time series data and could be viewed in-house. The driver face and steering wheel/hand position videos could only be viewed at the VTTI secure data enclave. A still cabin view was also available at the enclave, which showed a blurred view of the cabin that could be used to indicate passengers.

Data Request 1 (Phase 1)

The first data request was in Phase 1 and was made in the initial stages of the NDS and RID data collection. Only Florida was included in this request because collection of NDS and RID data was the most advanced in Florida at that time. Using the trip maps provided by VTTI, 50 rural two-lane curves



Florida

New York

Pennsylvania

North Carolina

Indiana

Source: World_Topo_Map (Esri, HERE, DeLorme, TomTom, Intermap, Increment P Corp., GEBCO, USGS, FAO, NPS, NRCAN, GeoBase, IGN, Kadaster NL, Ordnance Survey, Esri Japan, METI, Esri China [Hong Kong], swisstopo, MapmyIndia, © OpenStreetMap contributors, and the GIS User Community).

Figure 3.2. Location of identified segments by study state.

Table 3.3. Raw Data Output

System. Time	vtti.accel_x	vtti.accel_y	vtti.accel_z	vtti.pedal_ gas_position	vtti.gyro_y	vtti.gyro_x	vtti.wiper	vtti.gyro_z	speed
205	0.0116	–0.0087	–1.0063	12.54902	0	–0.3252		–0.3252	23.33335
206	0.0174	–0.0174	–0.9976	12.54902	0	–0.3252		–0.3252	23.33335
207	0.0203	–0.0058	–0.9947	12.54902	–0.3252	0		–0.3252	23.33335
208	0.0319	–0.0174	–1.0092	12.54902	0.325195	0		–0.3252	23.05557
209	0.0029	–0.0174	–0.9976	12.54902	0	–0.3252		–0.3252	22.7778
210	0.0261	–0.0029	–0.9918	12.54902	0	–0.65039		0	22.7778
211	0.0145	0.0029	–0.9947	12.54902					22.7778
212	0.0058	0.0029	–0.9976	12.81046	0	0		0	22.7778
213	0.0203	–0.0232	–0.9715	13.46406	–0.65039	0		–0.3252	22.7778
214	0.0029	–0.0232	–0.9831	13.92157	0	0		0	22.7778
215	0.0145	–0.0116	–0.9831	14.31373	0	0		–0.3252	22.7778
216	0.0145	–0.029	–1.0034	15.09804	0	–0.65039		–0.3252	22.7778
217	0.0232	–0.0203	–1.0005	15.55556	0.650391	–0.65039		–0.3252	22.7778
218	0.029	–0.0145	–0.9802	16.33987	–0.65039	0		–0.3252	22.7778
219	0.0174	–0.0116	–0.9715	16.60131	–0.97559	0		–0.3252	22.7778
220	0.0058	–0.0261	–1.0034	16.86275	0	0		–0.65039	22.7778
221	0.0261	–0.0261	–1.0063	17.12419					22.7778
222	0.0145	–0.0116	–1.0295	17.25491	0.650391	–0.3252		–0.65039	22.7778
223	0.0348	–0.0116	–0.9947	17.25491	0	0		–0.3252	22.7778
224	0.0377	–0.0232	–0.9686	17.25491	–0.65039	0		–0.3252	22.7778

in Florida were identified. Buffers representing geographic boundaries around each curve were developed and submitted to VTTI. After the initial data request was made (May 2012), it was determined that data were only available for eight buffer sections in Florida, which resulted in data for 14 curves because some buffers contain more than one curve.

Source: VTTI.

Figure 3.3. Example of video views.

The team requested that all vehicle activity through each identified buffer area be provided. The specific data elements from the DAS were included in the data request. The team worked with VTTI via e-mail and phone to refine the data request. Members from the CTRE/ISU and UI teams also visited the secure data enclave in Blacksburg, Virginia, in May 2012. This visit provided the opportunity to actually see the data and get a sense of the quality of data. This was particularly important for the UI team, whose members would later make a second visit to reduce data from driver videos.

The team received almost 400 traces. Traces were imported into a GIS program and reviewed, and some were removed from the data set because of the following issues:

- Potentially identifiable data were present.
- Drivers turned onto or off the roadway of interest within the curve.
- Construction zones were present.
- Traffic control was present within the curve (not identified earlier).
- Lane position was not available or was highly unreliable.
- Forward or face video data were missing (indicated by VTTI).

After removing traces with problematic data, a total of 137 initial usable traces through the various curves were identified. Researchers realized that requesting one forward roadway view for each segment would have allowed the team to identify locations where roadway conditions had changed or construction was present. This detail would have better guided the data request and was used in Phase 2.

Data Request 2 (Phase 2)

After data were collected and reduced for Phase 1, the data requests were refined. A sampling plan was developed based on the expected number of samples necessary for the statistical analyses, the time and resources available to reduce the data, and the ability of VTTI to provide the data in a timely manner given that the NDS data collection was concurrent with this research effort. Reduction of the driver face video at the VTTI secure data enclave was expected to be the limiting factor. It was originally estimated that about 1,000 traces could be reduced.

Given that this research project was one of the first applications of the NDS data, and based on the experience with the data in Phase 1, it was expected that unknown issues were likely to arise that would alter the sampling plan. As a result, two data requests to VTTI were planned for Phase 2. The first (Data Request 2) was for about 200 traces. The goal was to reduce these data, identify additional issues, and then use this information to make a more targeted final data request of about 800 traces (Data Request 3).

The 203 buffer segments identified for Phase 2 were provided in ArcGIS shape files and were provided to VTTI. When the second data request was completed (early November 2013) only 10% to 15% of the available NDS data was processed. Trips were found for about 80 of the buffers in this data request. One forward-view video for each of the 80 buffers was provided and reviewed by the CTRE/ISU team. Review of a single forward view for a roadway segment provided information such as presence of construction or other changes not evident in the RID or Google View. Issues were found with four buffers, and they were removed. VTTI identified about 1,455 traces across the remaining 76 buffers.

The CTRE/ISU team then worked with VTTI to set criteria for selection of about 200 of these traces for the second data request. The criteria are summarized in the following general terms:

- *Step 1.* Exclude traces in which the driver does not traverse at least 75% of buffer. In some cases, the driver turns onto or off the selected segment; these instances do not constitute through trips.
- *Step 2.* Exclude traces in which speed or lateral position were problematic or not working or in which GPS appears problematic.
- *Step 3.* Excluding traces identified in Step 1 or 2, select traces in which the following conditions are met:
  - Side or forward acceleration ≥0.3 g;
  - Speed ≥100 km/h (68 mph);
  - A crash/near crash had occurred; and
  - High alcohol readings were present.

  The intention of this step was to identify locations where a potential roadway departure had occurred or where other driver behaviors of interest were present.
- *Step 4.* From the remaining traces, select traces to balance age and gender for a total of 200 traces. When possible, include traces in which pedal position and steering wheel position are identifiable (both variables are used to determine when a driver reacts to the curve).

The VTTI and the CTRE/ISU and UI teams communicated back and forth to set filters for the above conditions. However, there was no easy method to identify when offset was not reliable. Unless only null values are present, it is not a simple task to determine that the lane-tracking system is not functioning properly or is producing erroneous data. The DAS does provide a variable for the probability that the lane-tracking system is correctly interpreting right- or left-side lane markings. However, no guidance is currently available regarding when the probability is low enough that the data should be discarded (e.g., probability ranges from 0 to 1,024, with higher values indicating better probability). As a result, it was difficult to know at what point to set the threshold. A threshold of 500 was set in Phase 1 to indicate reliable versus unreliable data and was refined to 512 in consultation with VTTI for Phase 2. Similar problems were present in determining when data such as speed or acceleration were valid.

Due to resource constraints, it was difficult for VTTI to check the data to determine how to better set filters and identify traces in which key output such as speed or offset were not available or reliable. It was decided that the most expeditious way to get data was for VTTI to provide the CTRE/ISU team with a data file for each of the 1,455 traces. Each spreadsheet contained DAS data such as position, offset, speed, and acceleration. The CTRE/ISU team reviewed all of the data and selected 200 traces for the second data request. Once the 200 traces of interest were identified, VTTI provided the forward and rear roadway videos.

Data Request 3 (Phase 2)

The team intended to use steering wheel position to indicate drowsy driving and to identify the point at which drivers began reacting to the curve. The research team also intended to ensure a subset of impaired drivers indicated by the alcohol sensor. However, steering wheel position data was much less available than expected because this variable could not be downloaded by the DAS in certain types of vehicles. Significant noise was also present in the alcohol measure, and it was not certain whether potentially alcohol-impaired drivers could be identified. As a result, presence of alcohol was not used as a filtering criterion. In addition, it was decided not to bias the sample toward vehicles for which steering wheel position data were available, because that might have resulted in oversampling of certain vehicle types.

By mid-March 2014, about one-third of the NDS data had been processed. VTTI queried the processed data using the provided buffers and identified an additional 2,647 traces. It was again determined that the most expeditious way to move the data request forward was for the CTRE/ISU team to review all of the available data and select a subset of about 800 traces. A Microsoft Excel macro was developed that summarized speed, lane position probability, side acceleration, and forward acceleration for each of the traces.

One forward-view video was also requested for each segment so that any unusual situations such as construction or recent changes to the roadway could be identified and those segments excluded.

Traces meeting the exclusion criteria in Step 1 or Step 2 were identified and removed from further consideration. About 60 of the viable traces met the criteria for side acceleration, forward acceleration, or speed in Step 3 and were selected.

The remaining viable traces were sorted into a matrix by curve and driver characteristics. An additional 720 traces were selected to balance curve and drivers characteristics, resulting in a total of 787 traces. VTTI provided the forward and rear roadway view for these traces. During later processing some additional issues were identified with the data files, resulting in some additional attrition.

Summary of Data Received and Limitations

Three separate data requests were made, as described in the previous section. About 137 traces were identified and reduced for Phase 1. In Phase 2, about 900 additional traces were determined to be viable after data screening and quality assurance were conducted, as described in the previous section. The data sets provided in-house for each trace included the following:

- One Excel file with GPS location and vehicle kinematic data;
- One forward video; and
- One rear video.

A total of 739 curves were initially identified and supplied in the full data request. Data were available for some curves, but trips did not traverse the entire segment. Additionally, only about one-third of the full NDS data set was available for query at the time the final data request was made. As a result, data were only available for 148 curves, which limited the number of curve characteristics that could be represented.

Although a large number of potential trips were ultimately available during Phase 2, there were a number of issues with the data (as is expected with this type of data collection). As a result, only a subset of traces was viable. The following issues with the some of the time series data were encountered:

- *Missing values.* In these cases, data such as speed are missing for all or portions of the trace.
- *Repeat values.* In these cases, the same value is repeated over multiple rows. This error is easy to identify because even travel at constant speed will produce minor fluctuations in values from row to row.
- *Erroneous values.* Values are reported incorrectly. This is usually evidenced by unusually high or low values. For instance, lateral acceleration is >0.3 g for 30 rows.
- *Irregularly reported values.* In most cases, variables such as speed, acceleration, lane offset, and pedal position are reported or averaged at 0.1-s intervals. In a number of cases, values were missing for a number of rows (e.g., reported every eighth interval). If the values are correct but less frequent, they can still be used in event-level analyses, but they are problematic when time series data are needed.
- *Data not available for all vehicles.* Because the DAS interfaces with the vehicle computer, some data, such as steering wheel position, could not be downloaded from all vehicles. Steering wheel reversal can be used to indicate traces in which drivers may have been drowsy and to indicate at what point a driver began reacting to the curve, but this information was only available for a fraction of vehicles.
- *Sensor accuracy unknown.* The accuracy of the head pose, lane offset, and alcohol sensor data had not been reported at the time data requests were made. Some indication of lane offset accuracy could be determined using lane line probability, plotting the data, and reviewing the forward view. Head pose could not be confirmed, and there was some indication that the alcohol sensor was not reporting consistently. As a result, neither the head pose nor alcohol data were included in any of the analyses.

Issues with the various video views include the following:

- Views are blurry due to glare and other factors.
- Views are missing.
- The driver's face or eyes cannot be seen because of glare, sunglasses, or other reasons.

Many of the variables were critical to the analyses, so it was important to screen out problem traces. Several attempts were made to set filters with VTTI so that problematic data were not included. However, as explained in the description

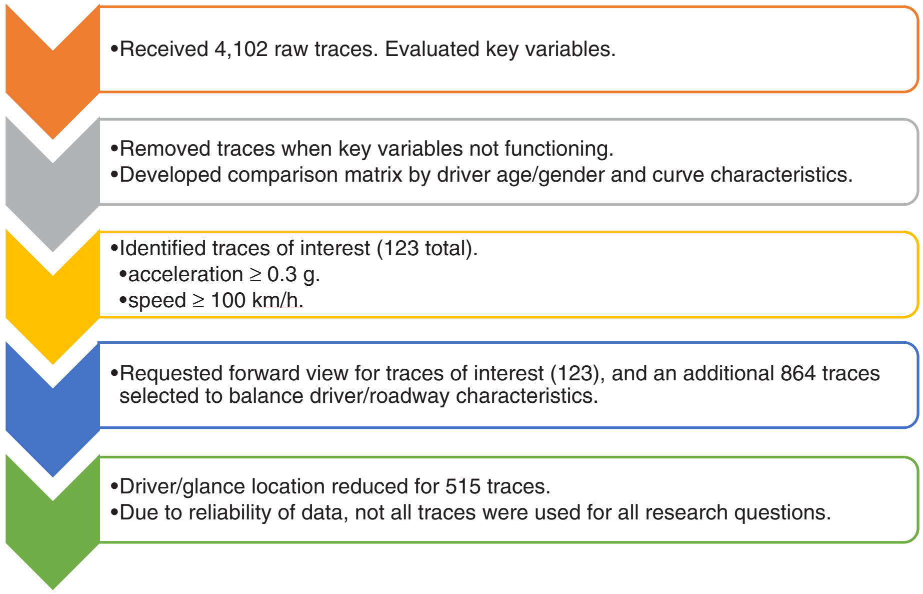


Figure 3.4. Data request and reduction process.

of data requests, the filtering process could not be fine-tuned as much as needed, so the CTRE/ISU team ended up reviewing a large number of traces that ultimately were not viable. This problem will likely have been addressed by the time final data quality assurance has been conducted by VTTI.

In some cases, a trace could be used for one research question but not others because necessary data were missing. For instance, pedal position was irregularly reported or missing in a number of traces. Consistent values were needed for Research Questions 1, 2, and 4 but not for Research Question 3. Research Questions 2 and 4 required consistent, accurate lane-tracking information, which further reduced the available traces. As a result, more traces were available for Research Question 3 than for the other research questions.

The data request and reduction process is shown in Figure 3.4.

CHAPTER 4

Data Reduction

This chapter describes how roadway, driver, and environmental data were reduced. Some additional data reduction may have been conducted for a specific research question and is described in the corresponding data section for that research question.

Reduction of Roadway Variables

Roadway variables were extracted for all of the curves identified for the data request, as described in Chapter 3, even though some curves ultimately did not have trips. The RID data collected as part of SHRP 2 Safety Project S04A was used to extract roadway variables when available. In some cases a variable was not collected, and in other cases the RID was not available for the study segment because the RID only covered a portion of the area where trips actually occurred. When the information was not available through the RID, other sources were used to manually extract the data. These additional sources were also used to confirm data collected through the RID, such as speed limit and advisory speed limit.

ArcGIS was used to measure distances between curves using the points of curvature included in the RID. ArcGIS was also used to determine whether the curve was an S-curve or a compound curve based on the distance between curves.

Google Earth was used to extract the roadway features not included in the RID. For instance, chevron presence was available for some of the states in the RID and was manually collected for those in which it was not. Radius was provided for most curves in the RID and was reported as radius by lane. When RID data were not available, radius was measured using aerial imagery. NDS forward video was used to determine subject measures for delineation, pavement condition, roadway lighting, and roadway furniture (which describes objects around the road that provide some measure of clutter), as well as a measure of sight distance for each curve. Variables collected are shown in Table 4.1.

The methodology for reducing roadway data is provided in Appendix A.

Given that the impact of countermeasures on roadway departure risk was a focus of this research, curves were selected to represent the widest range of countermeasures possible. Curve geometry is also highly relevant to roadway departure risk. Table 4.2 summarizes the number of curves with the indicated radius and countermeasures present.

Reduction of Vehicle, Traffic, and Environmental Variables

Each of the traces represents one driver trip through a selected roadway segment. One spreadsheet (containing DAS data), one forward video, and one rear-view video were provided by VTTI for each trace. Each row of data represents 0.1 s, and spatial location was provided at 1-s intervals. Several other variables reported at 1-s intervals include use of cruise control, air-bag deployment, date, and heading. A timestamp was also provided to link the various videos with the DAS data. A list of the main DAS variables provided includes the following:

- ABS activation
- Acceleration, *x*-axis
- Acceleration, *y*-axis
- Acceleration, *z*-axis
- Accelerator position
- Air bag, driver
- Alcohol
- Ambient light
- Cruise control
- De-identified date
- Dilution of precision, position
- Driver button flag
- Electronic stability control
- Elevation, GPS

Table 4.1. Roadway Variables Extracted and Main Source

Feature	ArcGIS	SHRP 2 RID	Google Earth	SHRP 2 NDS Forward Video
Curve radius		✓		
Distance between curves	✓			
Type of curve (isolated, S, compound)	✓			
Superelevation		✓		
Presence of rumble strips		✓		
Presence of chevrons		✓	✓	
Presence of W1-6 signs			✓	
Presence of paved shoulders		✓		
Presence of RPM			✓	
Presence of guardrail			✓	
Speed limit		✓		
Advisory sign speed limit		✓	✓	
Curve advisory sign/W1-6		✓	✓	✓
Pavement condition				✓
Delineation				✓
Roadway lighting				✓
Sight distance				✓
Roadway furniture				✓
Direction of curve				✓
Driveways along curve			✓	
Driveways along upstream section			✓	
Sight distance				✓

Table 4.2. Distribution of Curve Characteristics

	<1000	1000 to <1500	1500 to <2000	2000 to <2500	2500 to <3000	3000 to <4000	4000 to <6000	6000+	Total
Radius	31	19	23	19	14	15	13	14	148
Chevrons	5	3	0	0	0	0	0	0	8
Some paved shoulder	23	18	22	14	9	15	13	13	127
Rumble strips	0	0	0	0	1	1	0	0	2
RPM	4	4	8	5	2	5	4	1	33
Markings obscured or not present	6	3	0	1	1	0	1	0	12
W1-6	4	3	2	0	0	0	0	0	9
Lighting	0	0	0	0	0	0	0	0	0
Guardrail	5	6	5	4	3	0	7	1	31
On-pavement curve signing						Not present			
Flashing beacons or dynamic speed signing						Not present			

- Head confidence
- Head position *x*
- Head position *y*
- Head position *z*
- Headlight setting
- Lane marking, distance, left
- Lane marking, distance, right
- Lane marking, probability, left
- Lane marking, probability, right
- Lane marking, type, left
- Lane marking, type, right
- Lane position offset
- Lane width
- Pedal, brake
- Pitch rate, *y*-axis
- Radar range rate forward *x*
- Radar range rate forward *y*
- Roll rate, *x*-axis
- Seatbelt, driver
- Spatial position (lat/long)
- Speed, vehicle network
- Steering wheel position
- Time into trip
- Timestamp
- Wiper setting
- Yaw rate, *z*-axis

The static driver and vehicle characteristics available include the following:

- Driver
  - Driver age
  - Gender
  - Education level
  - Annual miles driven
  - Years of driving
  - Number of moving violations
  - Number of crashes
- Vehicle
  - Vehicle year
  - Vehicle model
  - Vehicle make
  - Vehicle track width

Smoothing

Vehicle variables were either available (i.e., acceleration, position, lane offset) or reduced (i.e., distance from right or left lane line) from DAS variables, which were provided at 10 Hz (one row = 0.1 s). A macro was developed that calculated lane position, change in pedal position, and change in steering position and that smoothed offset, lane position, pedal position, side acceleration, forward acceleration, and speed.

Smoothing was necessary because a certain amount of noise in the data resulted in improbable data points. Several different methods to smooth the data were investigated. The Kalman filter estimates the optimum average factor for each subsequent state using information from past states. It was determined that, although the Kalman filter was appropriate, developing a model for five different variables for more than 1,000 vehicle traces was overly complicated and time consuming.

A moving average method was selected because it is able to reduce random noise while retaining a sharp step response. Each of the variables listed above was smoothed using a moving average method, as follows (Smith 2003):

$$y[i] = 1/M \sum x[i+j]$$

where

$y[i]$ = the output signal;
M = the number of points used in the moving average; and
x = the input signal.

An example of smoothed versus original data is shown in Figure 4.1 for lane offset for a vehicle trace.

DAS Data Reduction

A macro was developed in Microsoft Excel to calculate additional columns in the DAS worksheets. They include the following:

- Spatial location in reference to each curve's point of curvature (PC) and point of tangency (PT) (e.g., 100 m upstream of the curve);
- Change in pedal position;
- Averaged pedal position (applied smoothing using a moving average method over five intervals);
- Forward acceleration (also calculated using the DAS accelerometer);
- Averaged forward acceleration (applied smoothing using a moving average method over five intervals);
- Change in steering wheel position (only available for a small portion of traces); and
- Averaged steering wheel position (applied smoothing using a moving average method over five intervals).

The macro also calculated distance from the right edge of the vehicle to the right edge line and distance from the left vehicle edge to the left lane line based on vehicle track width and offset.

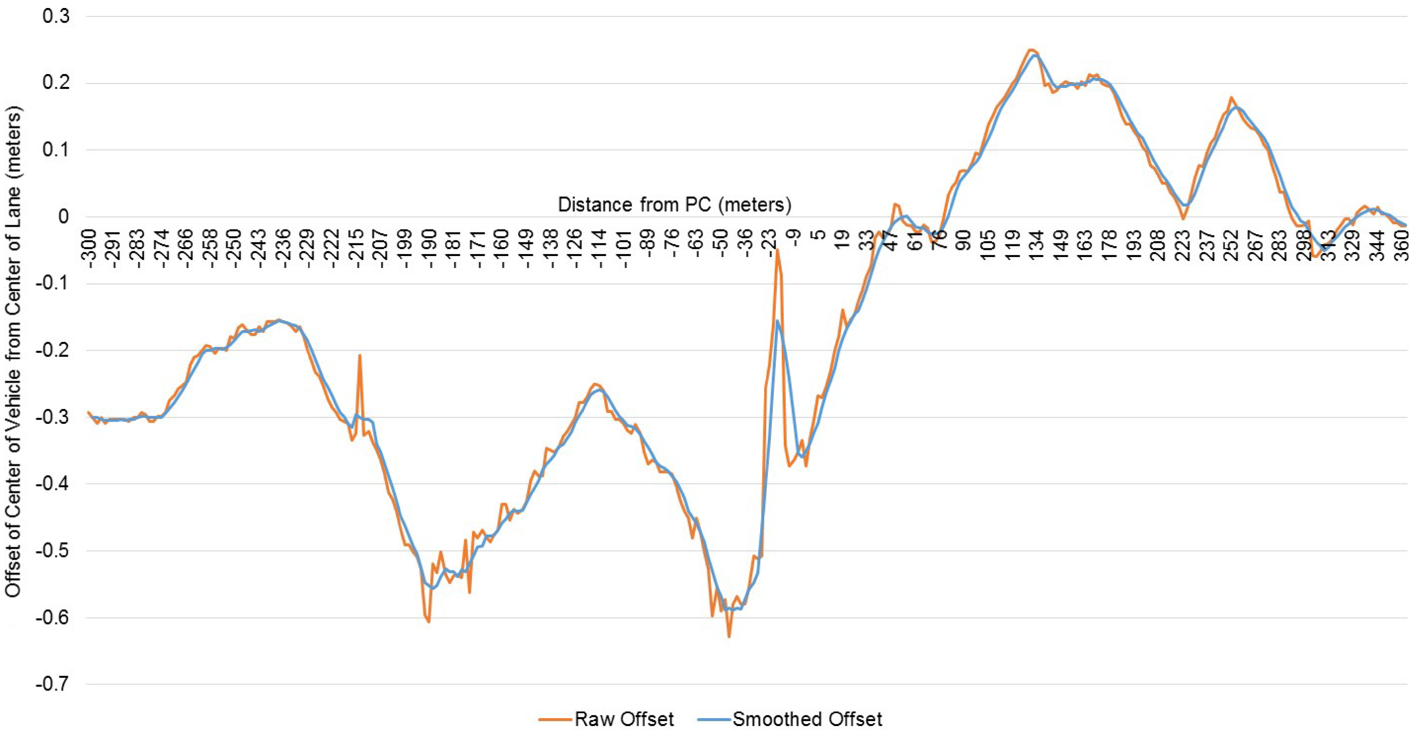


Figure 4.1. Original versus lane offset smoothed using moving average method.

Information on the make, model, class, and track width of each vehicle was also linked using the subject ID. In addition, the subject ID was used to link the driver demographics, such as the following:

- Gender;
- Age;
- Education;
- Work status;
- Household income;
- Number of miles driven last year;
- Average annual mileage for the last 5 years;
- Experience (i.e., number of years driving);
- Number and type of moving violations in the last 3 years;
- Number and type of crashes within the last 3 years; and
- Auto insurance for at least 6 months.

Vehicles traces were overlaid with the RID. For each curve, the nearest GPS points to the PC or PT were found and the position of the PC/PT located within the time series data using interpolation. Once PC/PT was established, vehicle position upstream or downstream of the curve was accomplished using speed. For some traces, there were multiple curves, so the PC/PT and upstream/downstream distances were determined for each curve. In some cases, speed was missing for multiple timestamps. In these cases, speed was interpolated assuming a constant increase or decrease.

Extraction of Data from Forward View

The forward video was used to reduce the environmental and other variables. The variables collected included the following:

- Surface conditions (e.g., dry, wet, snow);
- Lighting conditions (i.e., day, dawn, dusk, night with no lighting, night with lighting);
- Visibility;
- Locations of vehicles in the opposite direction passing the driver's vehicle;
- Locations where the driver's vehicle was following another car;
- Locations of curve advisory signs and locations where first visible;
- Locations of chevrons and locations where first visible; and
- Locations of potential roadway departure.

Reduction of Kinematic Driver Characteristics

Initially, three data collection trips were deemed sufficient to reduce 800 to 1,000 traces for Phase 2. However, because data collection trips at the secure data enclave were conducted at the same time VTTI was processing and conducting quality assurance on the NDS data, some issues were present that slowed reduction of the kinematic data significantly. A fourth trip was made, but only 515 total traces could be coded. These included one crash and three near crashes. This was significantly less data than planned and limited the amount of data that could be used in the different analyses.

Hawkeye, the VTTI-developed video data reduction tool that allowed analysts to simultaneously observe multiple camera views and use preset key strokes to code driver characteristics, was used to code data for this project.

Driver attention was measured by the location on which a driver was focused for each sampling interval. Scan position, or eye movement, has been used by several researchers to gather and process information about how drivers negotiate curves (Shinar et al. 1977). The majority of studies have used simulators to collect eye tracking information. Because eye tracking is not possible with NDS data, glance location was used as a proxy. Glance locations, shown in Figure 4.2, represent practical areas of glance locations for manual eyeglance data reduction. Note that Figure 4.2 does not show "over the shoulder," "missing," and "other" eyeglance locations. Those three locations were determined based on the UI team's extensive eyeglance reduction experience. Glance locations were coded using the camera view of the driver's face, with a focus on eye movements, but taking into consideration head tilt when necessary.

Potential distractions were determined by examining both the view of the driver's face and the view over the driver's right shoulder, which showed hands on/off the steering wheel. Distractions were identified when drivers took their eyes off the forward roadway. The coding process was developed by the UI team. They are experts in the field of human factors and have used similar methodologies in other NDS studies. Potential distractions include the following:

- Route planning (locating, viewing, or operating);
- Moving or dropped object in vehicle;
- Cell phone (locating, viewing, operating);
- IPod/MP3 (locating, viewing, operating);
- Personal hygiene;
- Passenger;
- Animal/insect in vehicle;
- In-vehicle controls;
- Drinking/eating; and
- Smoking.

Glance location and distractions were coded for each trace. The data reductionist indicated each time the glance location changed, and the data reduction tool recorded the timestamp. Similarly, the start and end times for distractions were also recorded. The data reduction method used to code driver glance location and distraction is provided in Appendix B.

Glance location and distractions were manually merged with the trace files using timestamp as a reference. Once this was completed, glance location was indicated for each row in the trace file. As a result, the time series analysis has glance location and distraction at the same resolution as the DAS variables.

A number of issues were noted during reduction of the driver face and steering wheel/hand position videos, based on the UI team's experience in reducing other data sets:

- Bright sunlight caused the camera to "wash out" the entire face, especially at certain times of the day when the sunlight

Figure 4.2. Glance locations.

was more direct. External light sources at night, such as street lights, created the same effect.

- Night videos had a grainy quality, making it difficult to distinguish facial features and almost impossible to code eye glances. For a large portion of the files, it was not possible to code glances without the use of head movement. Coding was especially difficult for glances within the cabin that require little head movement (e.g., console, steering wheel).
- Many drivers wear sunglasses, which completely obscure the eye, or prescription glasses, which create problems associated with glare.

Traces were used unless it was not possible to code glance location or distraction. It should be noted that glance and distraction were more likely to have been accurately coded for traces with clearer views of the face and eyes. However, discarding data that had some issues would have entailed removing almost all nighttime data and significantly reducing sample size.

Glance location was further reduced to indicate time spent with "eyes-off-roadway" engaged in roadway-related tasks or eyes-off-roadway engaged in non-roadway-related tasks based on data coding used by Angell et al. (2006). The authors define roadway-related glances or situation awareness (SA) as glances to any mirror or speedometer. Glances to other locations are defined as not roadway-related (NR).

Roadway-related glances (SA) included left mirror, steering wheel, and rearview mirror.

The data reductionists could not distinguish between a glance to the right mirror and a glance to the right for other reasons (e.g., to converse with passenger). Additionally, on a two-lane roadway, glances to the right mirror are not likely to be as common because drivers are not expecting vehicles to the right. Consequently, all glances to the right were considered to be non-roadway-related.

Additionally, when glances to roadway-related locations were also associated with a distraction, it was decided that these glances were likely to be non-roadway-related. For instance, a driver who was texting and glancing at the steering wheel was likely to be looking at the cell phone rather than the speedometer. As a result, non-roadway-related glances included center console, up, right, or down.

Data reductionists also indicated characteristics that applied to the trace in general, such as when the driver appeared to be drowsy or emotional. Weather conditions that add to the driving demand were also noted.

Summary of Data Limitations

About 1,000 traces were identified for Phases 1 and 2, and a data file with DAS variables, a forward-view video, and a rear-view video were provided in-house. Roadway, environmental, and static driver characteristics were reduced or provided for all of the available traces. As noted in this chapter, key DAS variables were not present or reliable for some traces, so not all traces could be used for all of the research questions.

Initially, three data collection trips were deemed sufficient to reduce 800 to 1,000 traces for Phase 2. However, because data collection trips to the secure data enclave were conducted at the same time VTTI was processing and conducting quality assurance on the NDS data, some issues were present that slowed reduction of the kinematic data significantly. A fourth trip was made, but only 515 traces total could be coded, which limited the amount of data that could be used in any of the analyses. Consequently, the main limitation to this study was a smaller than expected sample size.

Another limitation was that some types of data were not available and could not be included. Surface friction and pavement edge drop-off are important factors in roadway departure crash risk, but neither could reasonably be collected and were not available in the RID.

It would have been ideal to intentionally select a range of driver states, such as distraction or drowsiness, to ensure a reasonable sample of certain driver characteristics. However, there was no available method to detect whether driver distractions were present from the time series data so that traces with distraction could be preselected. Distraction could only be identified by viewing the driver face video, which was the last step in the data reduction process. Initially, the team identified a method to preselect traces in which drowsy driving may have occurred by using steering wheel reversals. However, steering wheel position data was only captured for a subset of vehicles. As a result, it was not possible to identify drowsy driving using the time series data. The team did target drowsy driving by intentionally including nighttime driving when available. Nighttime conditions were present for 124 traces (~25%), and 36 traces were at dawn/dusk. Additionally, because the accuracy of the alcohol sensor was not known at the time data were collected, potential impairment could not be targeted.

Another limitation was that newer vehicle technologies could not be targeted. Vehicles with electronic stability control (ESC) or collision warning systems made up only a small fraction of the vehicle fleet. Because other factors had a higher priority, vehicles with advanced technologies were not specifically targeted.

Curve characteristics were described earlier in his chapter. Other characteristics represented in the final available data set are described in Figures 4.3 and 4.4. The distribution of driver age and gender ($n = 202$) represented in the

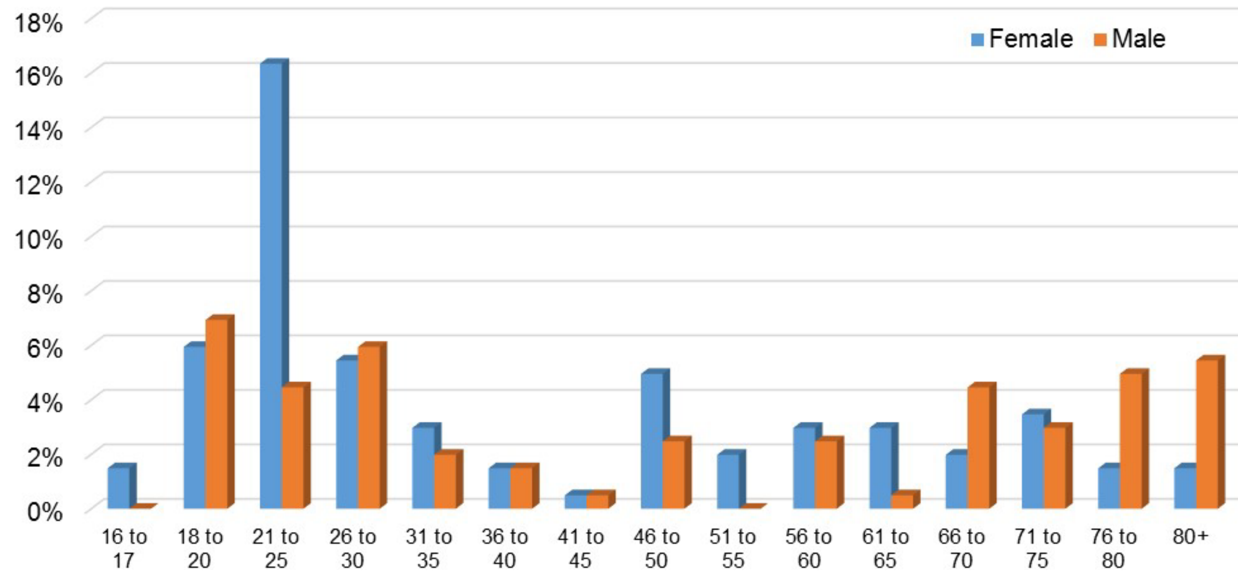


Figure 4.3. Distribution of driver age and gender.

viable traces is shown in Figure 4.3. A number of drivers had multiple trips.

Speeding was common, as indicated in Figure 4.4, which shows the percentage of drivers who entered the curve over the advisory speed limit if present or posted speed limit if not present by a certain threshold. Almost all drivers entered the curve 5 mph or more over the advisory speed, and a large fraction of drivers entered the curve 20 mph or more over the advisory speed. Data are summarized from the data reduced for Research Question 3 (sample size = 583).

A summary of the crash/near-crash events is provided in the sections below.

Summary of Crash 1

Event: 14950079
Driver: Male
Age: 21
Passengers: None
Location: WA

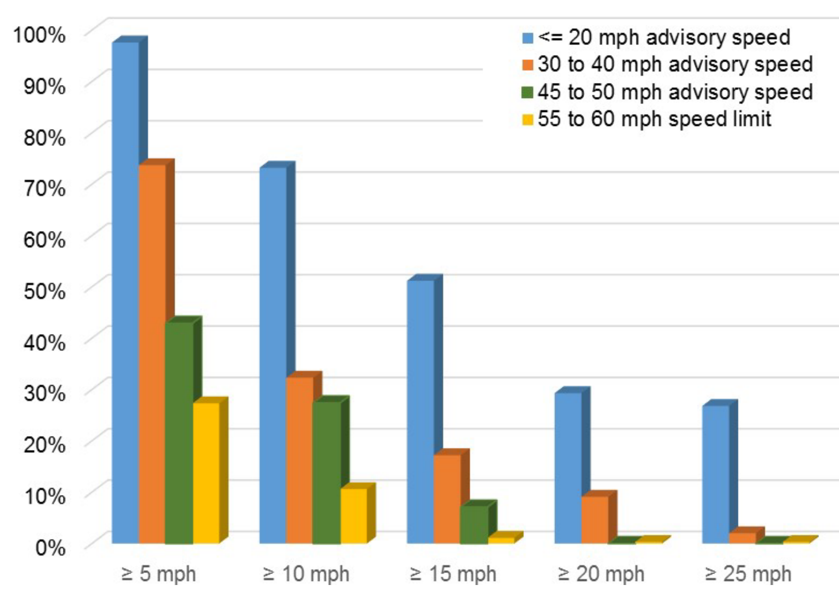


Figure 4.4. Percentage of drivers exceeding advisory or posted speed limit.

Month: August
Time of day: Midnight to 3:00 a.m.
Type of crash: Run-off-road
Description of site: This crash occurred on the second curve of an S-curve in rural Washington State. The driver was traveling from east to west. The curve on which the ROR crash occurred had a radius of approximately 50 m, with a lane width of approximately 3 m and no shoulder. This curve had guardrail on the inside portion of the curve and did have chevrons. The curve also had a curve advisory sign and a curve advisory speed of 20 mph, a reduction from the 45 mph speed limit on the rest of the roadway.

Data on seatbelt usage, ESC activation, and the status of traction control and cruise control were missing from the DAS data, but from the driver video it was determined that the driver was wearing a seatbelt. From the driver face video, it appears that the driver may have been tired, as he is looking forward and resting his head on his right arm/hand.
Description of crash: The driver took the right-hand curve too fast, departed the lane to the left at 4082869, and departed the roadway at 4084070 after crossing over opposing lanes of travel. The impact was with the ditch, bushes, and trees along side of roadway. The driver appeared to be unhurt.
Suspected main contributing factor: Speed appears to be the main contributing factor. The driver entered the curve traveling approximately 54 mph, 9 mph over the posted speed limit and 34 mph over the curve advisory speed limit. The only glance outside of a forward glance that occurred was a glance to the steering wheel, which lasted 0.6 seconds and occurred 7.7 seconds before the driver left his lane. This glance was not associated with a distraction.

Summary of Near Crash 1

Event: 11543161
Age: 18
Gender: Female
Passengers: Three (one front, two rear)
Location: VA
Month: July
Time of day: Noon to 3:00 p.m.
Type of crash: Near rear-end
Description of site: This near crash occurred on the tangent downstream of a curve. The roadway had lanes approximately 3.33 m wide, paved shoulders, centerline rumble strips, and raised pavement markings. The speed limit was 55 mph.

Data on seatbelt usage, ESC activation, and the status of traction control and cruise control were all missing from the DAS data. From the driver face video, it appears that the driver may have been tired, as she is looking forward and resting her head on her right arm/hand.
Description of near crash: The driver came up over a vertical curve while reaching to pick up a drink from the cup holder. The driver noticed that the vehicles ahead had come to a stop and had to brake hard to avoid a rear-end collision.
Suspected main contributing factor: The driver being distracted coupled with the presence of a vertical curve that limited sight distance appear to be the main contributing factors for this near crash. The driver was traveling under the 55 mph speed limit.

Summary of Near Crash 2

Event: 15483160
Age: 25
Gender: Female
Passengers: One front
Location: IN
Month: February
Time of day: 3:00 p.m. to 6:00 p.m.
Type of crash: Near rear-end
Description of site: This near crash occurred near the PC of a curve with a radius of 1,637 m. The speed limit on the roadway is 50 mph, with lanes approximately 3.2 m wide. There are paved shoulders and a curve advisory sign alerting drivers to the curve.

Data on seatbelt usage, ESC activation, and the status of traction control and cruise control were all missing from the DAS data. The driver is talking to a passenger throughout the trip.
Description of near crash: The driver looked away and was operating in-vehicle controls when the vehicles ahead came to a stop. The driver swerved to the right shoulder to avoid a rear-end collision.
Suspected main contributing factor: The driver is talking to the passenger during most of the event and that appears to be the main contributing factor for this near crash. The driver is operating the in-vehicle controls while looking at the center console from 1297157 to 1298625. The brake lights of the vehicle ahead are visible at 1297290. The participant vehicle does not begin braking until 1.3 s later. The driver was traveling at the 55 mph posted speed limit.

Summary of Near Crash 3

Event: 22512290
Age: 24
Gender: Male
Passengers: None
Location: NY
Month: August
Time of day: 3:00 p.m. to 6:00 p.m.
Type of crash: Near rear-end
Description of site: This near crash occurred near the PC of a curve with a radius of 570 m. The speed limit on the roadway

is 45 mph, with lanes approximately 3.5 m wide. There are paved shoulders and a curve advisory sign alerting drivers to the curve.

Data on seatbelt usage, ESC activation, and the status of traction control and cruise control were all missing from the DAS data. From the driver face video, it appears that the driver may have been tired, as he is looking forward and resting his head on his right arm/hand.

Description of near crash: The vehicle ahead applied its brakes at 392583. At this same time, the participant driver glanced at the center console. The driver braked hard to avoid hitting the vehicle ahead.

Suspected main contributing factor: The driver glancing away appears to be the main contributing factor to this near crash. The driver is looking at the center console from 391983 to 392850. The driver ahead slams on his/her brakes starting at 392583, so the participant driver does not notice the vehicle braking ahead for 0.3 seconds after the braking. The driver was only going about 1 mph over the speed limit of 45 mph, so speed did not appear to be a factor.

CHAPTER 5

Selection of Crash Surrogates

The goal of SHRP 2 and stakeholders interested in the outcomes of Project S08 research is understanding crashes, particularly severe and fatal crashes. For this reason, the best measure of analysis would be to study crash causes. However, even with the significant amount of data collected in the SHRP 2 NDS, crashes are rare events. At the time data requests were made, only one roadway departure crash and three near crashes had been identified. As a result, it was necessary to use crash surrogates for this project.

In addition, the use of crash data to address safety problems is a reactive approach, which is not able to take into account events that lead to successful outcomes. Consequently, the use of surrogates provides an opportunity to study what happens preceding and following an incident or event. For studying crash surrogates, the most significant advantage of naturalistic driving studies is that they provide a firsthand record of the events that precede crashes and incidents. Roadway, environmental, vehicle, and human factors can be extracted directly rather than from secondhand information, such as police records and crash databases, to identify relationships among factors that influence roadway departure crash risk. This firsthand information can also be used to determine the factors that lead to a positive outcome using crash surrogates.

The following sections discuss the rationale for selection of the identified crash surrogates.

Identification of Possible Roadway Departure Crash Surrogates

The team surveyed the literature on crash surrogates in general and crash surrogates that have been used specifically for roadway departures. Time to collision (TTC), also referred to as time to accident or time to conflict, is one of the most common crash surrogates (Burgett and Gunderson 2001; Gettman and Head 2003; Chin et al. 1992). The concept is logical and provides a repeatable and easily understood metric to assess level of crash risk. Risk can be measured as a function of TTC, where, at TTC = 0, the subject vehicle and another vehicle/object collide. This makes setting boundaries relatively straightforward.

However, to apply the concept of time to collision, the safety-critical event that results in a crash needs to be defined. For an intersection crash, this is a simple process because the safety-critical event is usually collision with another vehicle or pedestrian. The safety-critical event is not so easily defined for roadway departures because multiple crash outcomes could occur for a given roadway departure. For instance, the same roadway departure could result in a rollover or fixed object crash or, if the driver overcorrects, in the vehicle returning to the roadway and colliding with another vehicle. Because TTC depends on knowing the likely outcome, it is difficult to use TTC as a crash surrogate for roadway departures.

Use of TTC is also difficult because GPS data from the DAS are not accurate enough to locate the subject vehicle at a given point with sufficient precision to determine distances between objects. Initially, it was thought that calculation of TTC might be possible using distance from forward radar and the nearest strikable object. However, an initial review of the forward radar for several traces suggested that the radar output did not have sufficient detail to determine TTC with another vehicle or object. At the point data used in this study were obtained, the radar had not been processed. Once radar data have been processed, researchers may explore the possibility of using it for estimating time to collision.

Time to leaving the shoulder or distance intruded on the shoulder has been suggested as a measure of TTC (Dingus et al. 2008). Unpaved shoulder width is not collected with mobile data collection, and other methods to measure unpaved shoulder width are not sufficiently accurate to estimate time to leaving the shoulder. Distance intruded on the shoulder is related to lane deviation and will be included in this analysis, as described below.

Lane deviation is another measure used as a crash surrogate for both ROR crashes and crashes due to distraction (Donmez et al. 2006). Porter et al. (2004) used lateral placement and speed to evaluate the effectiveness of centerline rumble strips. Miaou (2001) developed a method to estimate roadside encroachment frequency and the probability distribution for the lateral extent of encroachments using an accident-based prediction model. Taylor et al. (2002) observed vehicle placement relative to the edge line using single versus double paint lines to delineate the presence of shoulder rumble strips. Hallmark et al. (2011) used lateral position to evaluate the effectiveness of edge line rumble stripes.

Description of Selected Crash Surrogates

The data necessary to identify each of the potential crash surrogates mentioned above was considered against what was available in the actual NDS. As already described, TTC was not feasible because the available data were not of sufficient accuracy to determine the distance of the vehicle from the nearest hazard.

Lane position or amount of encroachment was another surrogate measure used by researchers and ideally would have been used to address the research questions. Lane deviation is provided as "offset" in the DAS data. A number of other lane position variables are reported by the DAS that can be used to calculate other metrics, such as distance from the left or right lane line.

These variables include the following (shown in Figure 5.1):

- O = offset (distance from the vehicle centerline to the lane center, in cm).
- W = lane width (distance between the inside edge of the innermost lane marking to the left and right of the vehicle centerline, in cm).
- L_{CL} = distance from vehicle centerline to the inside of the left lane marking, in cm.
- R_{CL} = distance from vehicle centerline to the inside of the right lane marking, in cm.
- L_{PR} = probability that the lane marking evaluation is correct for the left-side lane line.
- R_{PR} = probability that the lane marking evaluation is correct for the right-side lane line.

Offset from lane center and distance from the right lane (R_D) or left lane (L_D) line are the metrics currently being used as crash surrogates. R_D and L_D are calculated as shown in the following equations (in meters).

$$L_D = -L_{CL} - T_w/2$$

$$R_D = R_{CL} - T_w/2$$

where

L_D = distance from left edge of vehicle to left edge of lane line;
R_D = distance from right edge of vehicle to right edge of lane line; and
T_w = vehicle track width.

Use of offset or lane position was explored as the main crash surrogate of interest for Research Questions 2, 3, and 4. At the time data were reduced for this research project, the accuracy of the DAS lane-tracking system had not yet been established. There was also no method by which the CTRE/ISU team could verify the accuracy of the reported offset and lane position values. As a result, these variables were examined for a number of traces, and several observations were made. First, there is a certain amount of noise in the various variables, as is to be expected from a large-scale data collection of this nature. As an example, lane position offset is shown in Figure 5.2 for one vehicle trace for a distance 300 m upstream and then through the curve. As noted, there is a significant amount of variation and several spikes that do not represent actual erratic changes in lane position. This was resolved in many cases by use of smoothing algorithms, as discussed in Chapter 4.

Second, the machine visioning algorithm depends on lane lines or differences in contrast between the roadway edge and shoulder to establish position. When discontinuities in lane lines occur, offset is reported with less accuracy (indicated as lane marking probability, which varies from 0 to 1,024, with higher values indicating better probability). Discontinuities occur for several reasons, such as lane lines being obscured,

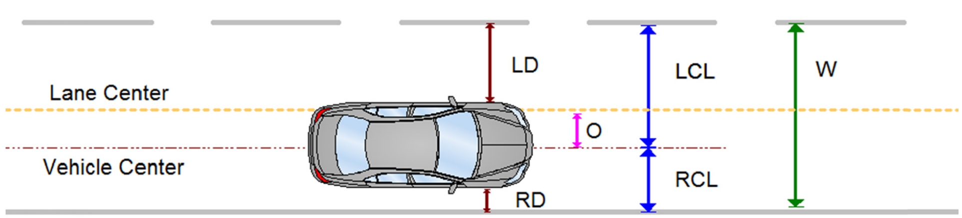


Figure 5.1. Description of lane position variables.

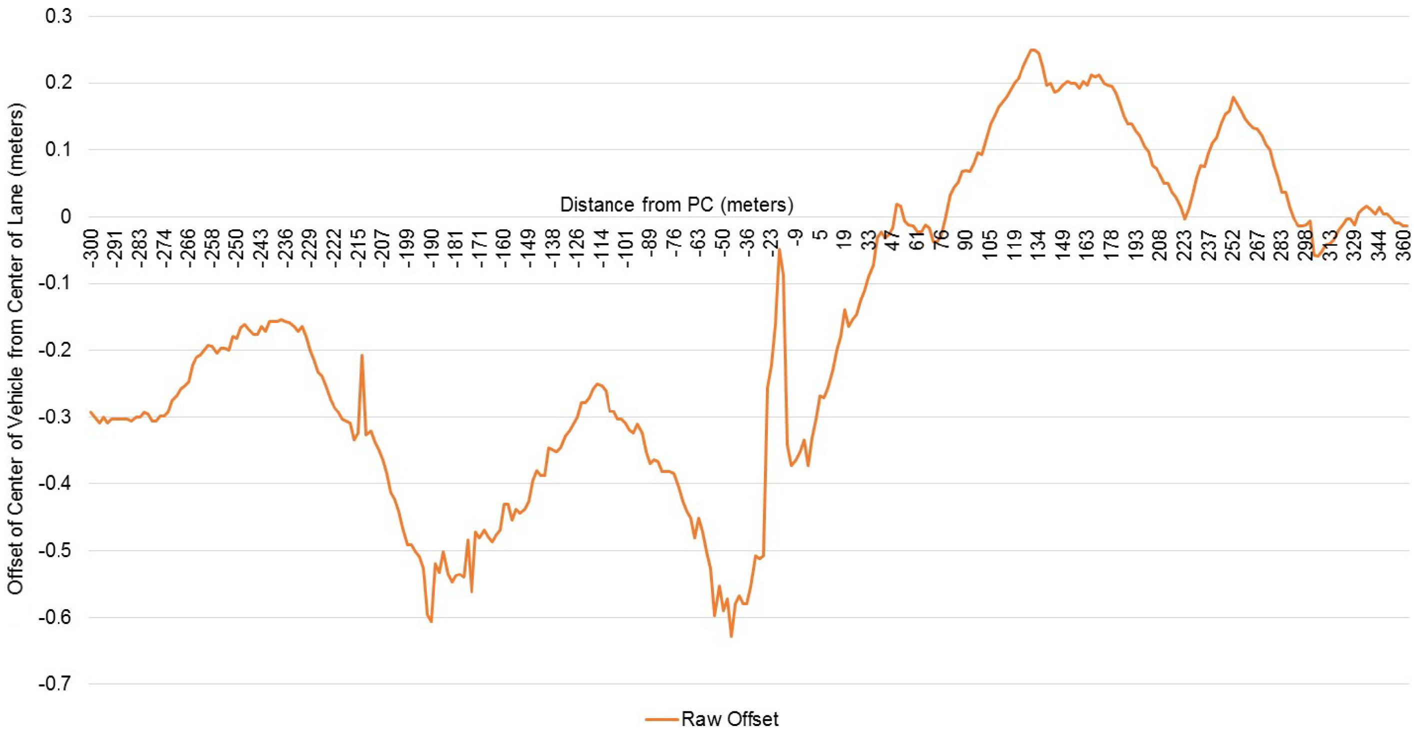


Figure 5.2. Offset in lane position.

natural breaks being present in lane lines (e.g., turn lanes, intersections), or visibility being compromised in the forward roadway view. An attempt was made to set a threshold to indicate the probability of reliable versus unreliable data. A threshold of 512 was selected based on review of the data and consultations with VTTI regarding their assessment of the lane-tracking system.

Third, in a number of cases the lane tracker did not appear to be working sufficiently to be considered reliable, but other indicators, such as side acceleration, suggested that a roadway departure may have occurred. Consequently, it was not realistic to exclude traces in which the lane-tracking system was not working.

As a result of these issues, lane offset or position could not be reliably used as a crash surrogate for a large portion of the data. Research Question 3 required the largest sample size in order to include a large number of driver, roadway, and environmental variables in the analyses. Logistic regression could also use binary dependent variables. As a result, it was determined that encroachments would be the best crash surrogate for Research Question 3.

A *right-side encroachment* is defined as the right vehicle edge crossing the right edge line (when present) or the estimated boundary between the lane and shoulder (when lane lines were not present). A *left-side encroachment* is defined as the left vehicle crossing the centerline. In all cases, the centerline was visible.

An encroachment was determined to have occurred when two of the following criteria were present:

- Vehicle edge is 0.2 m beyond edge line/centerline/lane–shoulder boundary.
- 0.2 g lateral acceleration is present.
- Edge line/centerline/lane–shoulder boundary crossing is visually confirmed using the forward view.

It should also be noted that left-side roadway departures may be drivers intentionally crossing the centerline (i.e., "cutting the curve"). However, it was not possible to identify when this occurred versus an inadvertent encroachment.

The amount of speed over the advisory or posted speed limit at curve entry was also used as a crash surrogate for Research Question 3. Although the correlation between speed and crashes on curves has not been established, speeding has been identified as a major crash contributor. Curve advisory and posted speed limit were known in all cases, and speed appeared to be universally present and reasonably accurate.

Research Questions 2 and 4 required a smaller sample size than Research Question 3. Additionally, offset was the only crash surrogate that made sense for the analyses selected. As a result, Research Questions 2 and 4 used offset or lane position as a crash surrogate and only included traces when offset or position were of sufficient reliability and continuity.

CHAPTER 6

Analysis for Research Question 1

This chapter discusses the first research question: What defines the curve area of influence?

Drivers begin to react to a curve at some distance upstream. This is expected to vary according to curve geometry, sight distance, and countermeasures present. Understanding where drivers begin to react to the curve is important for placement of traffic control and countermeasures. A better understanding of where drivers begin to react to a curve can help agencies determine the optimal placement of advance signing and other countermeasures.

Currently, placement of advisory signing along rural curves is primarily based on posted or 85th-percentile speeds and the amount of deceleration necessary for curve negotiation, following guidelines in the *Manual on Uniform Traffic Control Devices* (MUTCD), Chapter 2 (Federal Highway Administration 2009). When no deceleration is necessary, the distance varies from 31 m (100 ft) at 56 km/h (35 mph) to 122 m (400 ft) at 97 km/h (60 mph). Distances increase when speed reduction and lane changing in heavy traffic are expected.

Given that appropriate driver response upstream of a curve is necessary for proper speed selection and curve negotiation, defining the curve area of influence was necessary to determine how much data upstream of the curve should be included in the present analyses.

The objective of Research Question 1 was to identify where drivers begin reacting to a curve. A better understanding of where drivers begin to react to a curve can help agencies better determine placement of advance signing and other countermeasures. Research Question 1 was also used to indicate the curve area of influence for Research Questions 2, 3, and 4.

Data Sampling and Variables Used for Research Question 1

Use of eye tracking would have been ideal to determine where drivers were looking and noticing curves. However, eye tracking was not possible with the video data, and driver glance location could only be identified for general directions (e.g., left, right, steering wheel). As a result, glance location could not be pinpointed with sufficient accuracy to determine whether a driver noticed traffic control or roadway countermeasures. Therefore, vehicle kinematic data (i.e., braking or changes in speed) were the only method to assess at what point drivers began reacting to the curve.

Analyses in Phase 1 indicated that pedal position, speed, and steering wheel position could be used jointly to indicate the point at which a driver began react to the curve. Braking was only present in a few events and was therefore not used. After data were received for Phase 2, it became evident that steering wheel position was not universally recorded. As a result, change in speed and change in pedal position were the variables used to indicate where drivers began reacting to the curve.

Time series data were used for Research Question 1. Data may be output at different resolutions by the different sensors but are usually aggregated to 10 Hz (0.1-s intervals). Additional variables, such as vehicle position relative to the curve, were calculated and reported at the same resolution. Changes in pedal position and speed were smoothed over 0.5-s intervals and were calculated for each row to minimize the impact of noise using a moving average smoothing method. An example of time series data was shown in Table 3.2.

Speed was reported at 0.1-s intervals for the majority of the traces. Pedal position was also usually available but in many cases was reported at less frequent intervals (e.g., reported at 0.6- or 0.8-s intervals), which was too coarse for the models to detect changes. Consequently, only traces that had been reduced for the initial data request (about 200) that had both speed and pedal position reported at 0.1-s intervals were used in the analyses.

Additionally, only curves with a minimum distance of 400 m (1,312.3 ft) to the nearest upstream curve were used. Analyses in Phase 1 had suggested that drivers begin reacting to the curve within 200 m (656 ft), so provision of 400 m (1,312.3 ft) upstream allowed sufficient distance upstream of the expected reaction point to represent normal driving. Data were extracted for each curve of interest for a distance of 400 m (1,312.3 ft) upstream and through the curve.

Sample size was limited by the constraints described above. The analyses included 127 traces across 36 curves in Indiana, New York, and North Carolina. Curve radius varied from 117 m to 7,106 m (383.9 ft to 23,313.7 ft). Three curves had chevrons, four had W1-6 signs, none had rumble strips, five had guardrails, 21 had raised pavement markings, and 16 had curve advisory signs.

Methodology for Defining Curve Area of Influence

The point at which significant pedal position changes occurred was obvious in some traces, as shown in Figure 6.1, which shows pedal position for two traces (two drivers). In other cases, sufficient noise was present for pedal position and steering wheel position, so it was more difficult to identify the point of reaction, as shown in Figure 6.2.

A change point model was used to determine where drivers were reacting to the curve. A separate model was fit to each curve for each event using time series data. Change point models were fit using the statistical package R, which uses a regression model based on Muggeo (2008). The form of the model is as follows:

$$Y = \beta_0 + \beta_1 D + \beta_2 (D - D^*)$$

where

Y = the dependent variable for each model, which was either speed in meters per second or gas pedal position;

D = distance from the point of curvature, in meters; and

D^* = change point (the distance at which the driver reacts to the curve).

Note that distances are measured backwards from the point of curvature ($D = 0$), so all distances for this part of the analysis are negative. A change point model was selected because it could identify the point at which speed or pedal position changed significantly from upstream driving. The third parameter of the model, β_2, represents the strength of the reaction. If β_2 is not significantly different from zero, this indicates that the driver did not react in a noticeable way to the curve.

Thus, for this model the researchers were most interested in the values of β_2 and D^*, because D^* indicates the point at which the driver reacted to the curve and β_2 indicates how strong that reaction was. The identified reaction points are only meaningful if the strength of the reaction, the value of β_2, is significantly different from zero. So, the estimated β_2 value for each model was tested against the hypothesis that $\beta_2 = 0$.

Models were developed independently for speed and pedal position for each curve, for travel in the direction inside of the curve and outside of the curve. The reaction distance was then compared to MUTCD sign placement values.

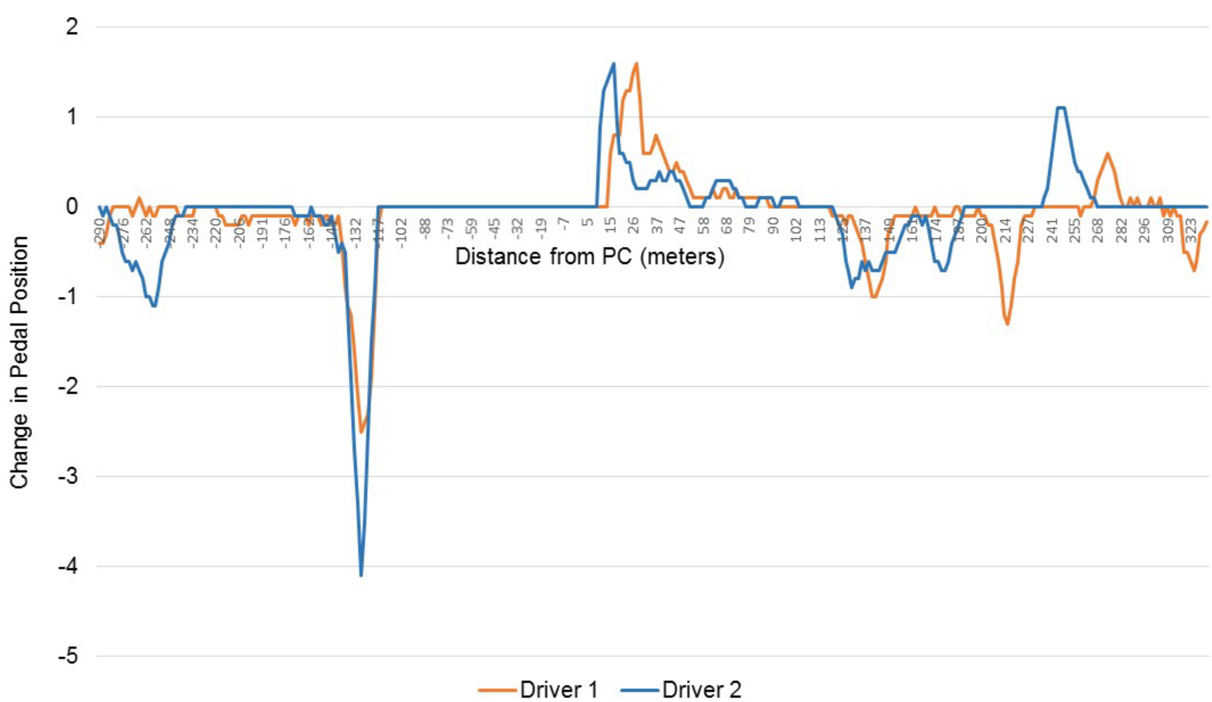


Figure 6.1. Change in pedal position for two drivers (Florida Curve 101).

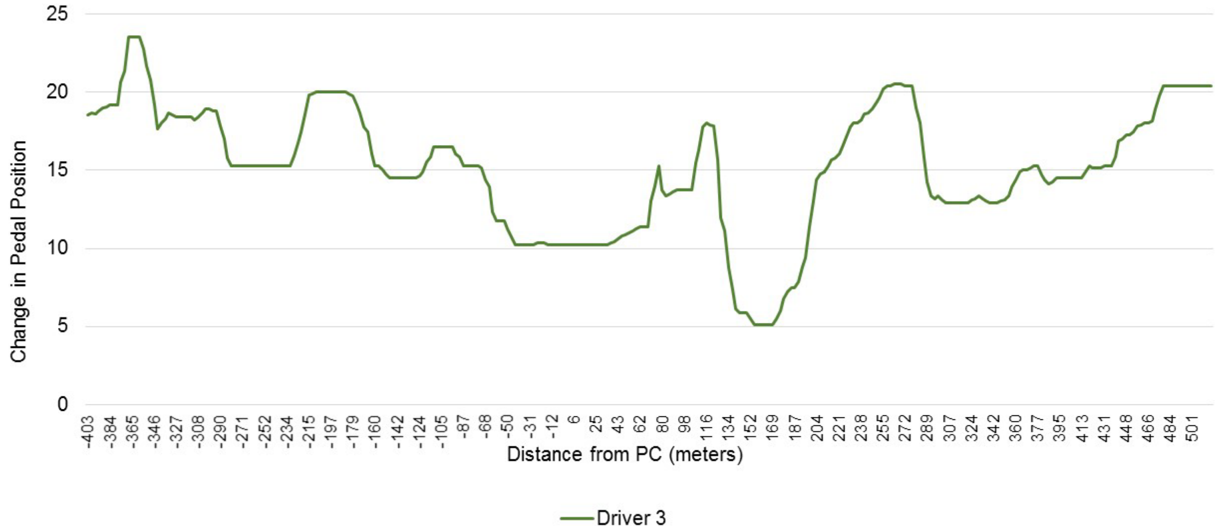


Figure 6.2. Change in pedal position for one trace in which change is not obvious.

A second model was developed for traces for 11 curves in Indiana for which multiple drivers were available. A Bayesian hierarchical change point model is as follows:

$$Y_{ij} = \beta_{0_{ij}} + \beta_{1_{ij}} D + \beta_{2_{ij}} \left(D - D_{ij}^{*}\right) + \beta_3 R_i + \beta_4 C_i$$

$$\beta_{k_{ij}} \sim Normal\left(\beta_{k_i}, \tau_k^2\right)\left(\begin{array}{c}\text{accounts for variability among}\\ \text{drivers, for } k = 0, 1, 2\end{array}\right)$$

$$\beta_{k_i} \sim Normal\left(\beta_k, \phi_k^2\right)\left(\begin{array}{c}\text{accounts for variability between}\\ \text{curves, for } k = 0, 1, 2, 3, 4\end{array}\right)$$

$$\beta_k \sim Normal(0, 100)(\text{noninformative prior})$$

$$\tau_k^2 \sim Inverse\ Gamma(0.01, 0.01)\left(\begin{array}{c}\text{noninformative}\\ \text{hierarchical prior}\end{array}\right)$$

$$\phi_k^2 \sim Inverse\ Gamma(0.01, 0.01)\left(\begin{array}{c}\text{noninformative}\\ \text{hierarchical prior}\end{array}\right)$$

where

i indexes the curve;

j indexes the driver; and

R_i and C_i are the radius and travel direction of curve i, respectively.

Results for Research Question 1

Speed Model

The fitted change point models with speed in meters per second as the dependent variable are plotted for each curve, as shown in Figures 6.3 to 6.5. A model was developed for each trace for each curve. Model results are provided in Appendix C.

Results are shown graphically in Figures 6.3 to 6.5 and are grouped by state and by curve and direction of travel (inside versus outside of the curve). Traces for many curves have similar reaction points, as indicated by the slope of the line changing at about the same point, such as curves IN44Ain and IN44Jout. Others, however, have distinctly different reaction points, as shown for curves IN13Ain and IN13Aout.

The average point at which drivers reacted for all curves was 164 m (538.1 ft) upstream of the curve. Results were averaged by curve radius, as shown in Table 6.1. When the models were tested for significant reactions, 96 of the 127 models were found to have significant driver reactions at the 95% confidence level, as shown in Appendix C.

Pedal Position Model

Change point models were also developed using pedal position as an indicator of driver response. This variable has no units, but is a measure of how far the driver is pushing down on the gas pedal. If the value increases, the driver is increasing pressure on the pedal, and if the value decreases, the driver is decreasing pressure (letting up) on the pedal. For the same 127 traces across 36 curves in Indiana, New York, and North Carolina, the fitted change point models show pedal position as the dependent variable. Model results are provided individually in Appendix C. Results are plotted graphically in Figures 6.6 to 6.8 and are grouped by state, curve, and travel direction (inside versus outside of the curve).

As noted, some of the curves have very similar reaction points for all events, which can be seen in curves IN13Aout and IN44Iout. Others, however, have very separated reaction points (e.g., curves IN44Ain and IN44Dout).

Again, the values of β_2 and D^{*} are of the most interest. The furthest reaction point was approximately 488 m (1,601.1 ft)

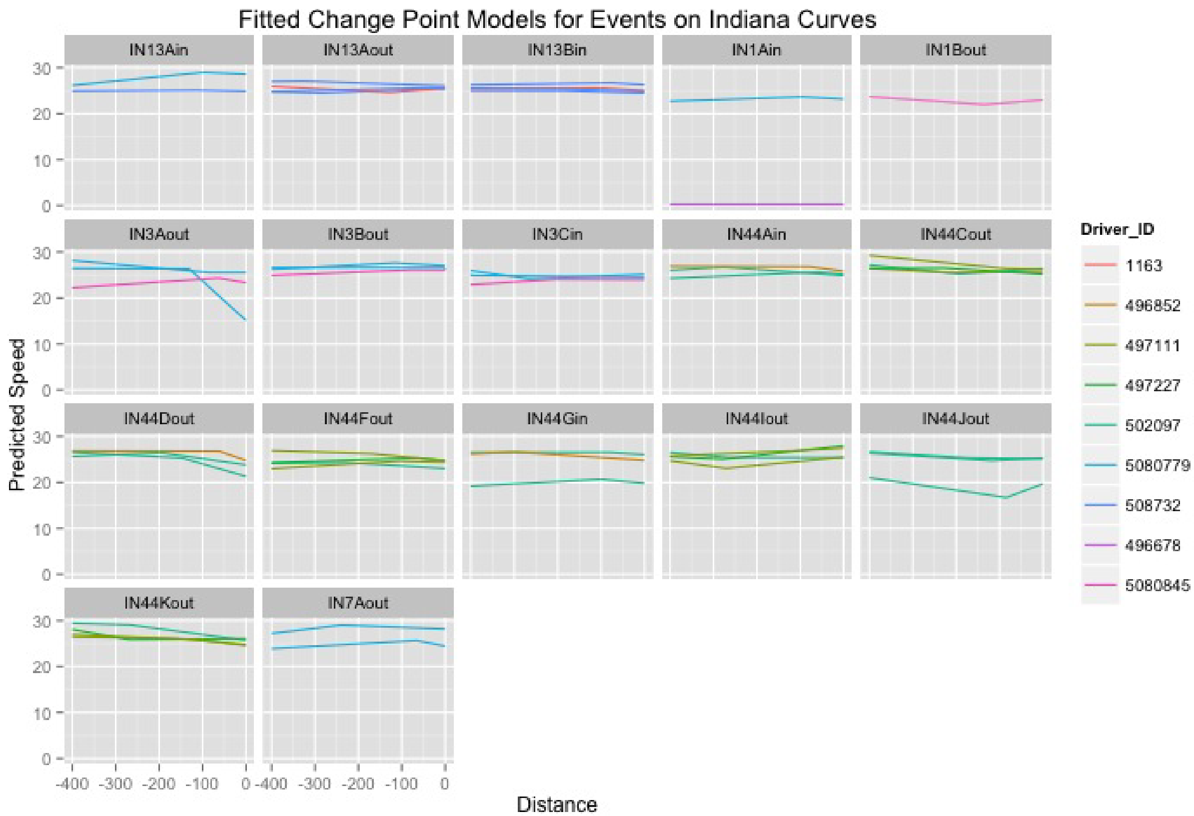


Figure 6.3. Fitted speed models for Indiana curves.

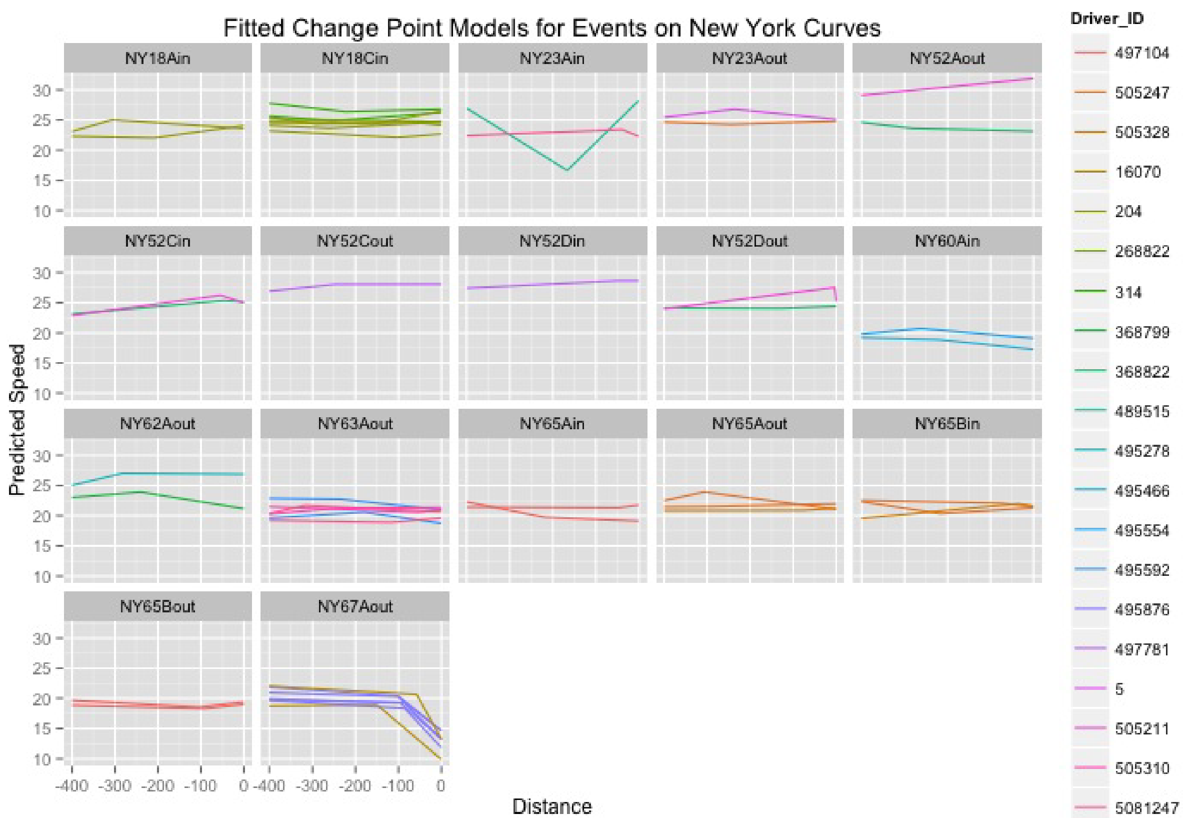


Figure 6.4. Fitted speed models for New York curves.

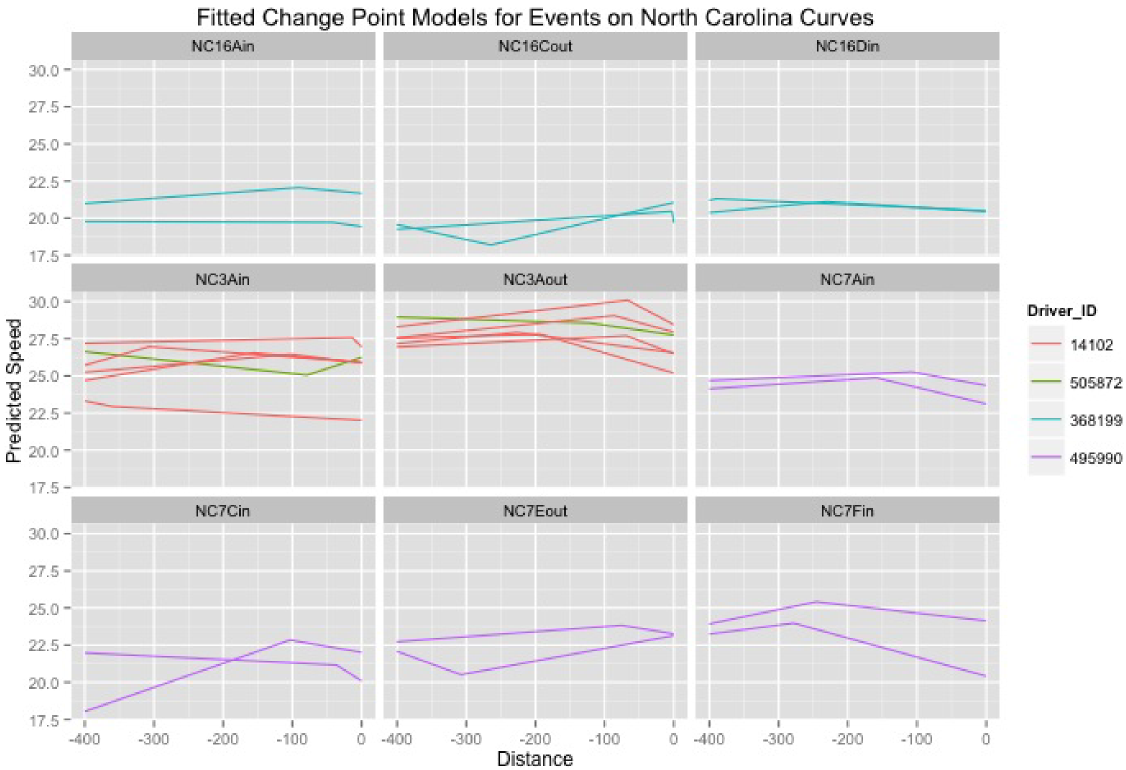


Figure 6.5. Fitted speed models for North Carolina curves.

before the point of curvature, and the closest reaction point was approximately 13 m (42.7 ft) before the point of curvature, with the mean reaction point about 179 m (587.3 ft) before the point of curvature. The estimated β_2 value for each model was again tested against the hypothesis that $\beta_2 = 0$, and 99 of the 127 models were found to have significant driver reactions at the 95% confidence level.

The average point at which drivers reacted for all curves was 180 m upstream of the curve, which was similar to the results for the speed models. Results were averaged by curve radius, as shown in Table 6.2. Results for individual models are shown provided in Appendix C.

Table 6.1. Speed Change Points Results by Curve Radius

Radius in Meters (feet)	Average Change Point in Meters (feet)	Number of Curves
<1000 (3281)	−142.9 (−468.8)	4
1000 to <1500 (3281 to <4921)	−146.1 (−479.3)	7
1500 to <2000 (4921 to <6562)	−193.2 (−633.9)	11
2000 to <2500 (6562 to <8202)	−191.1 (−627.0)	6
≥2500 (≥8202)	−149.8 (−539.7)	8

Results for Bayesian Model

The Bayesian model was only fit to a subset of the data: 11 curves in Indiana for which there was adequate repetition of drivers across curves. This model is an improvement on the current model because it is able to account for the individual differences among drivers, as well as the differences between curves. This model also allows the prediction of appropriate reaction points for other curves not included in the study, which could aid in deciding where to place chevrons and/or dynamic speed feedback display units to lower crash incidents on rural curves.

After accounting for the radius of the curve, the travel direction of the curve, and the variability among drivers, all curves were found to have about the same reaction point, approximately 105 m (344 ft) upstream of the curve. The 95% posterior credible interval for this estimate is from 136 m to 64 m (446 ft to 210 ft) upstream of the curve.

The exact reaction point for each curve changes with curve radius, and curve direction is given by the estimates of β_3 and β_4. The estimate of β_3 is −0.000872, with a posterior 95% credible interval of (−0.00143, −0.000284). So, for every additional 100 m (328.1 ft) in the radius of the curve, the reaction point moves back from the point of curvature by 0.0872 m (0.29 ft). The estimate of β_4 is −1.991, with a posterior 95% credible interval of (−3.112, −0.73). The curve directions

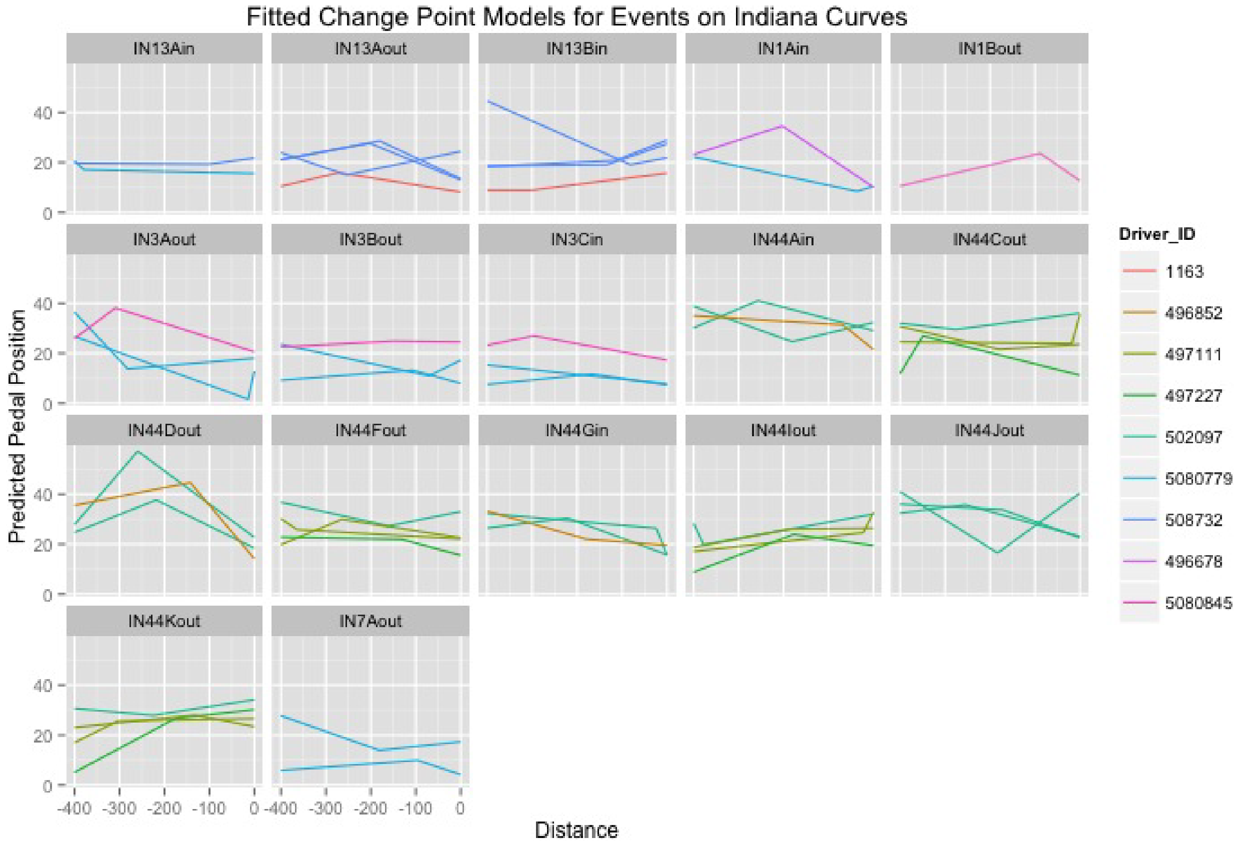


Figure 6.6. Fitted pedal position models for Indiana curves.

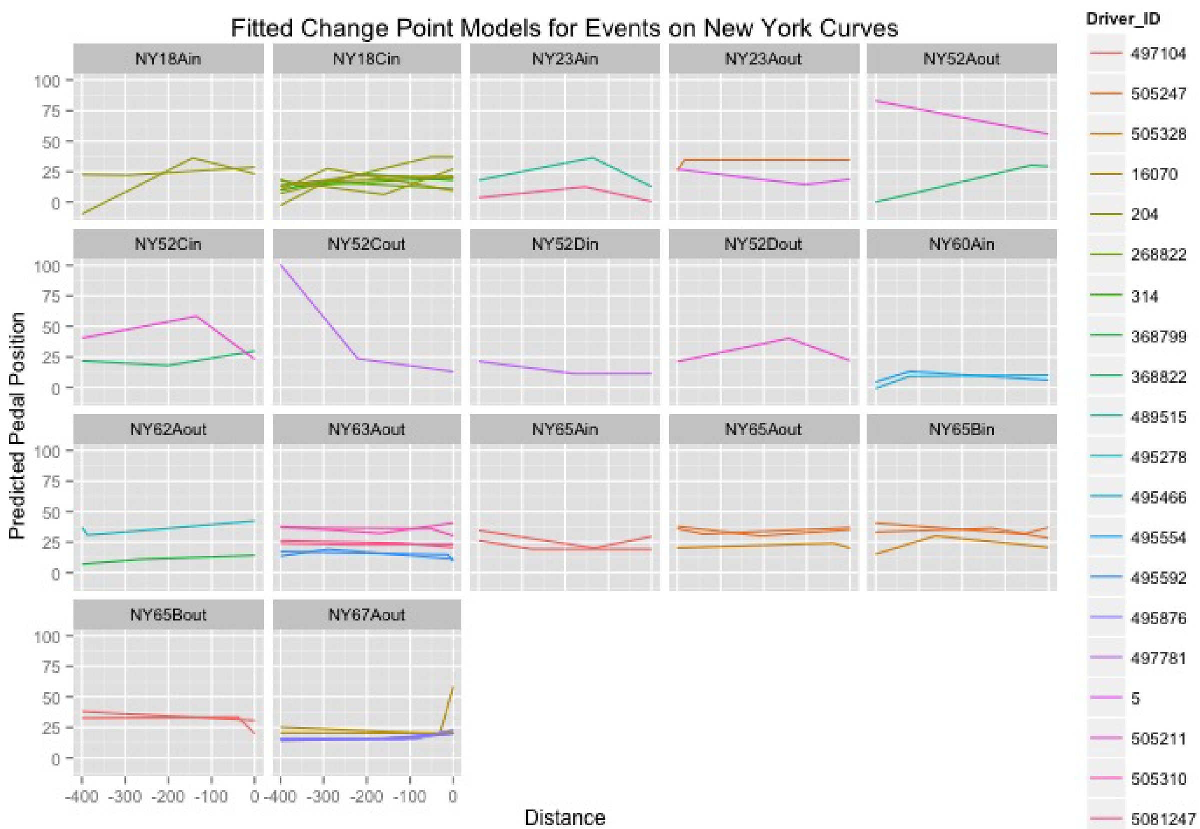


Figure 6.7. Fitted pedal position models for New York curves.

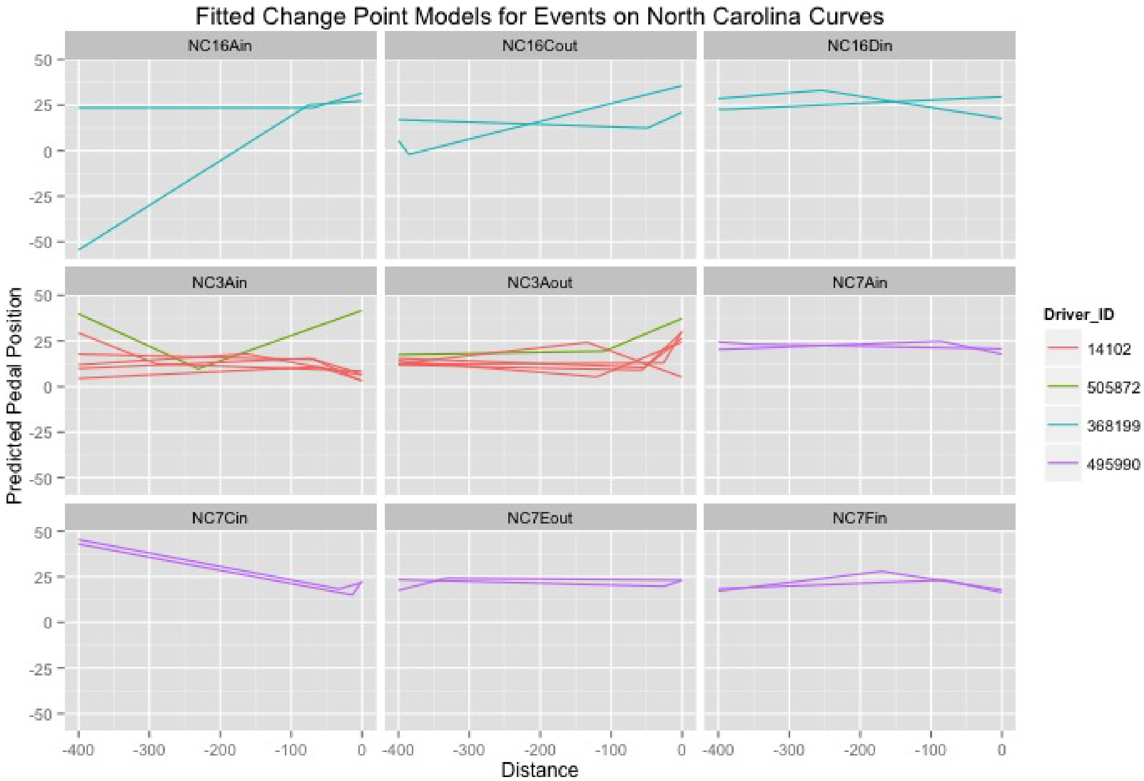


Figure 6.8. Fitted pedal position models for North Carolina curves.

were coded as 0 for inside and 1 for outside, so this estimate indicates that the reaction point moves on average about 2 m (6.6 ft) further from the point of curvature when the driver is traveling on the outside of the curve.

Summary and Discussion

The objective of Research Question 1 was to identify the point at which drivers begin reacting to a curve. A better understanding of where drivers begin to react to a curve can help agencies better determine placement of advance signing and other countermeasures. Research Question 1 was also used to indicate the curve area of influence for Research Questions 2, 3, and 4.

Key Findings

Time series data were modeled using regression and Bayesian analysis. The point at which speed and pedal change are significantly different from that of upstream driving was identified. Results indicate that, depending on the radius of the curve, drivers begin reacting to the curve 164 m to 180 m (538.1 ft to 590.6 ft) upstream of the point of curvature.

Results did suggest that drivers begin reacting to the curve sooner for curves with larger radii than for curves with smaller radii, as shown in Table 6.3. This was unexpected because sharper curves are more likely to have advance signing, chevrons, or other countermeasures that have the express purpose

Table 6.2. Pedal Position Change Points Results by Curve Radius

Radius in Meters (feet)	Average Change Point in Meters (feet)	Number of Curves
<1000 (3281)	–137.4 (–450.8)	4
1000 to <1500 (3281 to <4921)	–163.9 (–537.7)	7
1500 to <2000 (4921 to <6562)	–198.1 (–649.9)	11
2000 to <2500 (6562 to <8202)	–205.6 (–674.5)	6
≥2500 (≥8202)	–186.0 (–610.2)	8

Table 6.3. Average Change Point

Radius in Meters (feet)	Average Change Point in Meters (feet)	
	Pedal Position	Speed
<1000 (3281)	–137.4 (–450.8)	–142.9 (–468.8)
1000 to <1500 (3281 to <4921)	–163.9 (–537.7)	–146.1 (–479.3)
1500 to <2000 (4921 to <6562)	–198.1 (–649.9)	–193.2 (–633.9)
2000 to <2500 (6562 to <8202)	–205.6 (–674.5)	–191.1 (–627.0)
≥2500 (≥8202)	–186.0 (–610.2)	–149.8 (–539.7)

of getting a driver's attention. Additionally, drivers traveling at appropriate speeds do not need to reduce speed to the same extent on curves without advisory speeds as for curves where deceleration is necessary.

There may be several reasons for the unexpected results. First, countermeasures that simply warn drivers of an upcoming curve may not be sufficient to change driver behavior. Better delineation of the curve may be more effective in providing the appropriate roadway cues. Most of the curves with smaller radii had some type of advance warning, but only three had chevrons, which are highly visible in all environmental conditions. Three additional curves had raised pavement markings (RPMs), but RPMs are not as obvious during daytime conditions as during nighttime and wet weather conditions. Due to the sample size available, it was not possible to draw relationships between reaction point and presence of a specific countermeasure.

It is possible that driver reaction for sharper curves is more pronounced, and as a result, the models were better able to identify the reaction point. Another explanation for the unexpected results is that drivers may indeed be reacting to advance signing and delineation and are more gradually slowing than for curves with larger radii. Sight distance may also be an issue for sharper curves.

Implications for Countermeasures

The MUTCD (Federal Highway Administration 2009) suggests placement of warning signs based on posted or 85th-percentile speed, the driver's ability to decelerate to the posted advisory speed for the condition, and an assumed legibility distance of 76 m (250 ft). Sign placement for posted/85th-percentile speeds from 72 km/h to 89 km/h (45 mph to 55 mph) range from 31 m to 99 m (100 ft to 325 ft). As a result, the point at which a driver is able to view a sign (assuming favorable visibility and sight distance) is 107 m to 175 m (350 ft to 575 ft).

It should be noted that driver reaction point may be influenced by signs, and as a result, some correlation exists between presence of signs and reaction point. However, it was assumed that a warning sign only provides information to the driver and does not in and of itself cause the driver to react sooner.

The average point at which drivers begin reacting to the curve is summarized by curve radius in Table 6.3. This represents the reaction point for drivers who successfully negotiated the curve. Given that warning signs are only likely to be present for curves with smaller radii, sign placement falls within the reaction distance, suggesting that sign placement distances are appropriately set.

The results showing that drivers react sooner to curves with larger radii indicates that advisory signs and advisory speeds may not be sufficient to alert drivers to the upcoming curves. Countermeasures that provide better curve delineation, such as chevrons, may provide better cues to drivers so that they can gauge the sharpness and respond appropriately.

Limitations

One of the major limitations of this analysis is that driver glance location could not be used to detect driver response to an upcoming curve. Braking and steering may have provided additional insight but were not sufficiently available to include. As a result, the models depended on change in speed and pedal position to detect reaction point.

The major limitation to these speed and pedal position analyses is that, even with smoothing, there was a significant amount of noise. As a result, it was difficult to detect reaction point.

Sample size was also a limitation in this analysis. The sample size was limited by the number of traces with reliable pedal position values that were available in the data that could be reduced within the project constraints. If additional data were included, the models might be able to relate reaction point to countermeasures.

CHAPTER 7

Analysis for Research Question 2

This chapter discusses the second research question: What defines normal curve negotiation?

A methodology proposed by several other researchers was used to develop a relationship between tangent speed and offset and curve speed and offset, which was used to define normal curve driving. This relationship assumes that there is some relationship between the tangent and curve speed (drivers who speed upstream are also likely to speed within the curve) and between lateral offset upstream and within the curve (drivers who do not maintain lane position upstream will have similar lane keeping within the curve).

Schurr et al. (2002) developed a relationship between the operating speed on tangent sections 183 m upstream and at the curve midpoint. Stodart and Donnell (2008) collected data upstream and within six curves using instrumented vehicles with 16 research participants. They compared change in speed and lateral position from the upstream tangent to the curve midpoint using the following:

$$\Delta V = V_{\text{MPT}} - V_{\text{MC}}$$

$$\Delta L = L_{\text{MPT}} - L_{\text{MC}}$$

where

V = speed;
L = lateral position;
MPT = midpoint of tangent; and
MC = midpoint of curve.

The objective of Research Question 2 was to define normal curve driving. Understanding how a driver normally negotiates a curve during various situations provides insight into not only how characteristics of the roadway, driver, and environment potentially influence how a driver drives, but also the areas that can lead to roadway departures. Knowing how much drivers normally deviate in their lane, as well as how they choose their speed, could potentially have implications on policy or design.

Data Sampling and Segmentation Approach for Research Question 2

The conceptual model of curve driving assesses changes in metrics as the driver negotiates the curve based on the factors of the curve and driving behavior upstream of the curve. Data for several positions along the curve were sampled from the time series data from the DAS, which also had additional variables such as driver characteristics and environmental characteristics.

The sampling plan for the curve model can be seen in Figure 7.1. Data were sampled at each point shown (e.g., PC); the locations for sampling were determined after plotting events and determining which sampling scheme picked up the common patterns identified. Sampling in the tangent section was based on distance. Sampling within the curve was at equidistant points rather than at a specified distance because the curves have varying lengths.

The points sampled within the curve were the PC, PT, and then four equally spaced points (C1, C2, C3, and C4, as shown in Figure 7.1). Upstream data were collected every 50 m up to 300 m. These locations were chosen to capture driving upstream of where drivers react to the curve (i.e., normal tangent driving) along with the area where they react and as they approach the entry to the curve. Because the data sampling plan required 300 m of upstream data, the analysis included only isolated curves (i.e., no S-curves or compound curves) and curves with a tangent section that was at least 300 m from the nearest upstream curve.

DAS variables were sampled at each point shown. The variables included offset; speed; environmental characteristics, such as whether there was an oncoming vehicle or whether the driver was following another vehicle; and driver kinematic data, such as glance and distraction. Data collected for the upstream area included the offset and speed at each sample point, along with driver glance location and distractions.

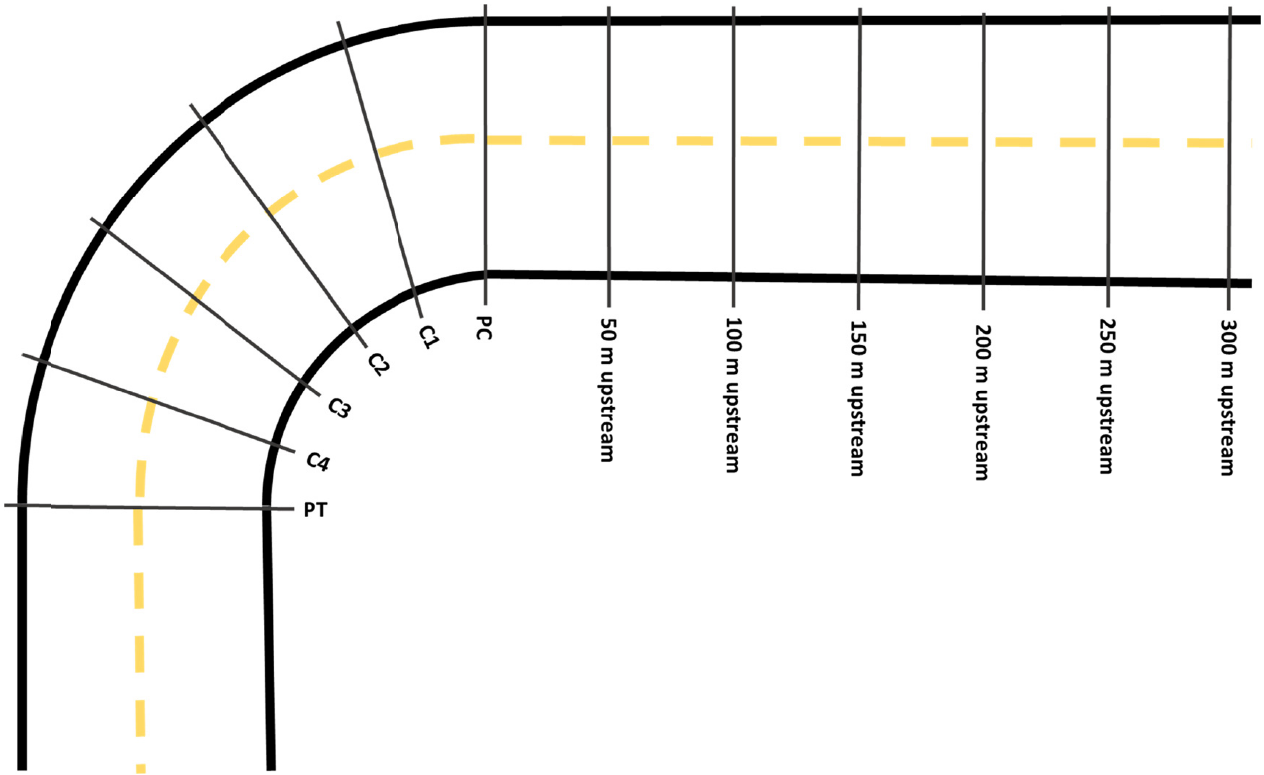


Figure 7.1. Data sampling layout for curve driving model.

In addition, as noted previously, a moving average method was used to smooth data across rows (0.5-s intervals) to reduce noise present in the variables. These data were merged with environmental and driver data. The complete list of variables collected is included in Chapter 4.

Because vehicle offset was the crash surrogate used for this research question, the offset data had to be quite accurate, as small changes in the offset could drastically affect the results of the model. Additionally, only isolated curves with a distance of 400 m or more to the nearest upstream curve were used.

Time series data for curves that met this criteria were examined, and data were only used when the team had a high level of confidence in the reliability of the lane offset variable for the entire tangent and curve sections that were sampled, as shown in Figure 7.1. Data were ultimately available for 17 curves. Thirty-six traces were available for the inside (right-hand curve) model, and 26 were available for the outside (left-hand curve) model.

Drivers were distributed by age and gender, as shown in Table 7.1.

Variables Used for Research Question 2

The conceptual model evaluates changes in driver attention and response expressed as changes in vehicle kinematics to model curve driving. The independent variables used in the models were (1) offset (in meters) and (2) the amount the driver was driving over the speed limit or curve advisory speed if present (in mile per hour). The dependent variables examined, along with a description of each, can be seen in Table 7.2.

Description of Analytical Approach for Research Question 2

Models for lane position and speed were developed for both inside (right-hand curve from the perspective of the driver) and outside (left-hand curve from the perspective of the driver).

For the lane position models, a generalized least squares (GLS) model was used. A panel data model was tested with

Table 7.1. Driver Characteristics for Research Question 2

	Age			Total
	16 to 25	26 to 50	>50	
Inside curve (right-hand)				
Male	0	4	3	7
Female	4	1	3	8
Outside curve (right-hand)				
Male	0	2	2	4
Female	5	1	5	11

Table 7.2. Variables Used for Research Question 2

Variable	Description
Curve PT	Factored variable that indicates the position in the curve from which data are sampled (PC, C1, C2, C3, C4, or PT)
Direction	Direction of the curve (0: outside/left, 1: inside/right)
Radius	Radius of the curve (m)
Chevrons	Indicator variable for chevrons (0: not present, 1: present)
Rumblestrips	Indicator variable for rumble strips (0: not present, 1: present)
Guardrail	Indicator variable for guardrail (0: not present, 1: present)
RPM	Indicator variable for raised pavement markings (0: not present, 1: present)
AdvisSign	Indicator variable for curve advisory sign (0: not present, 1: present)
Nighttime indicator	Indicator variable for nighttime (0: daytime or dawn/dusk, 1: nighttime)
SpeedUp	Speed limit in upstream (mph)
SpeedDiff	Speed differential between tangent and curve advisory speed (mph)
Over300	Amount over the speed limit at 300 m upstream of curve (mph)
Over250	Amount over the speed limit at 250 m upstream of curve (mph)
Over200	Amount over the speed limit at 200 m upstream of curve (mph)
Over150	Amount over the speed limit at 150 m upstream of curve (mph)
Over100	Amount over the speed limit at 100 m upstream of curve (mph)
Over50	Amount over the speed limit at 50 m upstream of curve (mph)
Overspeed	Amount over the speed limit at point in curve (mph)
Speed (mph)	Speed at point in the curve (mph)
Offset	Distance offset from centerline in points throughout curve (m)
Offset300	Distance offset from centerline 300 m upstream of curve (m)
Offset250	Distance offset from centerline 250 m upstream of curve (m)
Offset200	Distance offset from centerline 200 m upstream of curve (m)
Offset150	Distance offset from centerline 150 m upstream of curve (m)
Offset100	Distance offset from centerline 100 m upstream of curve (m)
Offset50	Distance offset from centerline 50 m upstream of curve (m)
OffsetSD	Standard deviation of offset in 300 m upstream of curve (m)
Distracted200	Visual distraction in 200 m upstream of curve indicator (0: not present, 1: present)
Distracted150	Visual distraction in 150 m upstream of curve indicator (0: not present, 1: present)
Distracted100	Visual distraction in 100 m upstream of curve indicator (0: not present, 1: present)
Distracted50	Visual distraction in 50 m upstream of curve indicator (0: not present, 1: present)
Distracted	Visual distraction in curve indicator (1: distraction present, 0: no distraction)
Forward	Forward glance at point in curve indicator (1: glance is forward, 0: glance away)
SA	Roadway-related glance (1: roadway-related glance, 0: otherwise)
NR	Non-roadway-related glance at point in curve indicator (1: present, 0: not present)
NRup	Non-roadway-related glance in 200 m upstream of curve indicator (1: present, 0: not present)
NRcurve	Non-roadway-related glance in curve indicator (1: present, 0: not present)
Visibility	Visibility indicator (1: low visibility, 0: otherwise)
Surface	Surface condition (0: dry, 1: pavement wet but not currently raining, 2: snow present, but roadway is bare)
PaveCond	Pavement condition (0: normal surface condition, 1: moderate damage, 2: severe damage)

(continued on next page)

Table 7.2. Variables Used for Research Question 2 (continued)

Variable	Description
Delineation	Delineation condition (0: highly visible, 1: visible, 2: obscured)
Shoulder	Paved shoulder width (1: less than 1 ft, 2: 1 ft to less than 2 ft, 3: 2 ft to less than 4 ft, 4: greater than or equal to 4 ft)
LargeShoulder	Paved shoulder greater than or equal to 4-ft indicator (0: not present, 1: present)
Gender	Gender indicator (0: female, 1: male)
Under25	Age under 25 indicator (0: over 25, 1: under 25)
Under30	Age under 30 indicator (0: over 30, 1: under 30)
Age	Age of driver at time of first drive
LargeVeh	Large vehicle (i.e., truck or SUV) indicator (0: car, 1: truck or SUV)

"EventID" as the individual and "Point in Curve" as the time setting. The Breusch-Pagan Lagrange multiplier test found that no panel effect was present; therefore, an ordinary least squares (OLS) model was appropriate. After running the OLS models, it was determined that there were problems with autocorrelation due to the spatial nature of the data, so a GLS model was used to correct for this.

The GLS function in the NLME package of R was used to develop the models. Models were selected to minimize Akaike information criterion (AIC) and Bayesian information criterion (BIC), while including significant variables from the list in Table 7.2. The correlation between the dependent variables and independent variables and the correlation between independent variables were examined to determine which variables should potentially be included in the model. The order of autoregression parameter was tested using an analysis of variance (ANOVA) test. The correlation structure of the model took into account the grouping across each event through each unique curve. The grouping factor allows for the correlation structure to be assumed to apply only to observations within the same unique event.

For the amount-over-the-speed-limit model, an OLS regression model was developed. The dependent variable was the amount over the speed limit (or curve advisory speed) at point C2. Modeling the amount over the speed limit for just this point was chosen as opposed to modeling speed over the entire curve because the only significant difference in speed was found at the PT, which is not of interest in the context of speed's role in lane-departure crashes in curves.

The model was developed using the lm() function of the R software package while attempting to have the best fit. Variables were included if they were significant at the 95% confidence level. Two outlier observations were not included because they skewed the fit of the model. Diagnostic tests showed that the assumptions of normality, linearity, independence, and homogeneity were met.

Results for Research Question 2

Four models were developed; lane position and amount of the speed limit were used as dependent variables; and models were developed for both inside (right-hand curve) and outside (left-hand curve) driving, because drivers tend to behave differently in each direction of curve. The dependent variable for lane position was offset of the center of the vehicle from the center of the travel lane. Positive offset is to the right, and negative offset is to the left of the center of the lane.

Second-order autoregressive GLS models were developed for both lane position models. Panel models were developed for both speed models. The results of the lane position and speeding models are discussed below.

Lane Position Model

The best fit model for lane position for right (inside) curves was developed using 216 observations and contained eight variables. The list of variables and parameter estimates is shown in Table 7.3. The model suggests that as drivers tend to the right (toward the edge line) in the upstream, the offset in the curve will also shift to the right, or near the outside of the lane.

If the driver is engaged in an eyes-off-roadway distraction at a particular point in the curve, the driver's lane position is expected to shift to the right near the outside of the lane 0.14 m at the next point. For instance, if a driver is engaged in an eyes-off-roadway distraction at 50 m upstream, the driver's lane position is expected to shift right 0.14 m at the PC.

The model also suggests that for every year older a driver is, the driver's lane position is expected to move toward the right 0.00345 m.

Finally, the model includes indicator variables relating to the position in the curve. At position C1 (see Figure 7.1),

Table 7.3. Significant Variables for Right Curve Lane Position Model

Variable	Parameter Estimate	*p*-Value
Constant	–0.22185	0.0005
Offset at 100 ft upstream of curve	0.36714	0.0000
Distracted at the previous point in the curve or upstream indicator (0: not distracted, 1: distracted)	0.13592	0.0500
Driver's age (years)	0.00345	0.0001
C1 position indicator (0: not C1, 1: C1)	0.16931	0.0001
C2 position indicator (0: not C2, 1: C2)	0.18865	0.0012
C3 position indicator (0: not C3, 1: C3)	0.39609	0.0000
C4 position indicator (0: not C4, 1: C4)	0.26790	0.0000
PT position indicator (0: not PT, 1: PT)	0.17682	0.0020
First-order autoregression disturbance parameter (phi 1)	0.57808	
Second-order autoregression disturbance parameter (phi 2)	–0.28316	
Number of observations	216	

which is just past the point of curvature, the average position is 0.17 m to the right of the center of the lane. At position C2 the average position is 0.19 m. As the driver gets closer to the center of the curve (position C3), the average lane position is 0.40 m to the right, which is a shift of more than 0.2 m from the upstream position. Drivers then move back toward the center of the lane at positions C4 and the PT (0.27 m and 0.18 m, respectively). As indicated, a driver's drift to the outside lane edge near the center of the curve suggests that the driver may be most vulnerable to a right-side roadway departure near the center of the curve.

These parameters support the idea that drivers do not maintain a smooth path through the curve. The first-order autoregression parameter phi was found to be 0.57808, and the second-order was –0.28316.

The best fit model for lane position for left (outside) curves was developed using 156 observations and included nine variables, as shown in Table 7.4. The parameter for offset at 100 m is similar to that in the right curve lane position model, just at a slightly greater magnitude. The model suggests that if a driver tends to drive to the right of the lane center upstream of the curve, the driver will also tend to drive to the right of the lane center within the curve.

When drivers engage in a non-roadway-related glance at a particular location, their lane position is expected to move to the left, or toward the centerline, by 0.13 m.

At night, lane position shifts toward the left (toward the centerline) by 0.12 m, which could potentially occur because there are fewer oncoming vehicles. When a large paved shoulder (≥4 ft) is present, the model predicts that the driver will move toward the right (toward the edge line) by 0.21 m; this is expected because the driver has more space than when no paved shoulder is present.

Indicator parameters for position in the curve were also included. While the parameters for indicators C4 and PT were not significant, they were still included because they give some information on the change in position throughout the curve.

As drivers enter the curve and move to the center of the curve (positions C1 to C3, as shown in Figure 7.1), they tend to be positioned about 0.16 m to 0.21 m to the left of the center of the lane (toward the centerline). As drivers move to the end of the center of the curve (position C4 and the PT), they move to the right toward the center of the lane. This suggests that drivers may be most likely to cross the roadway centerline as they enter the curve.

Speed Model

The amount a driver was over the advisory speed if present or posted tangent speed if not present was modeled at point C2 using OLS. The best fit model had an adjusted R^2 value of 0.741 ($n = 60$) and five variables. Speed model variables are shown in Table 7.5.

Model results show that for every 1.6 km/h (1.0 mph) over the speed limit a driver is traveling at 100 m upstream of the curve, the amount over the speed limit at point C2 is expected to increase by 1.1 km/h (0.7 mph). This result is expected

Table 7.4. Significant Variables for Left Curve Lane Position Model

Variable	Parameter Estimate	*p*-Value
Constant	0.00067	0.9904
Offset at 100 ft upstream of curve	0.44811	0.0002
Nonroadway glance at point in curve indicator (0: roadway glance, 1: nonroadway glance)	–0.13466	0.0193
Night indicator (0: daytime, 1: night)	–0.12283	0.0155
Paved shoulder greater than 4 ft indicator (0: paved shoulder less than 4 ft, 1: paved shoulder ≥ 4 ft)	0.21273	0.0006
C1 position indicator (0: not C1, 1: C1)	–0.17691	0.0014
C2 position indicator (0: not C2, 1: C2)	–0.20758	0.0044
C3 position indicator (0: not C3, 1: C3)	–0.16169	0.0304
C4 position indicator (0: not C4, 1: C4)	–0.02272	0.7495
PT position indicator (0: not PT, 1: PT)	0.05718	0.4071
First-order autoregression disturbance parameter (phi 1)	0.49063	
Second-order autoregression disturbance parameter (phi 2)	–0.26283	
Number of observations	156	

because drivers who are traveling over the speed limit in the tangent are also likely to speed within the curve. Speeds are expected to be 2.8 km/h (1.7 mph) slower at nighttime than during the day.

Additionally, the model suggests that those who drive a truck or SUV are expected drive 2.1 km/h (1.3 mph) faster than those who drive a passenger vehicle. The model also found that for every additional 10 years in age for a driver, speed decreases by 0.9 km/h (0.5 mph). Finally, the model suggests that if drivers are engaged in a non-roadway-related glance at point C2, they are expected to be driving 5.3 km/h (3.3 mph) slower than if they were glancing at the roadway.

Summary and Implications

The objective of Research Question 2 was to define normal curve driving. Understanding how a driver normally negotiates a curve during various situations provides insight into not only how characteristics of the roadway, driver, and environment potentially influence how a driver drives, but also the areas that

Table 7.5. Significant Variables for Speed Model

Variable	Parameter Estimate	*p*-Value
Constant	3.70299	<0.001
Amount over the speed limit at 100 m upstream of curve	0.70772	<0.001
Driver's age (years)	–0.05340	<0.001
Night indicator (0: daytime or dusk, 1: nighttime)	–1.73462	0.028
Vehicle type (0: car, 1: truck or SUV)	1.30152	0.029
Non-roadway-related glance at current point indicator (0: roadway-related glance, 1: non-roadway-related glance)	–3.32218	0.037
Number of observations	60	
Adjusted R^2	0.741	

can lead to roadway departures. Knowing how much drivers normally deviate in their lane, as well as how they choose their speed, could potentially have implications on policy or design.

Conceptual models of curve driving were developed to assess changes in lane position and speed as the driver negotiates the curve. Understanding how a driver normally negotiates a curve during various situations provides insight not only into how characteristics of the roadway, driver and environment potentially influence driving behavior, but also into areas that can lead to roadway departures.

Additionally, the model indicates boundaries for normal driving. Originally, the intent of answering this research question was to use this information to identify events of interest (nonnormal driving) to help establish boundaries between noncrash roadway departure events for Research Question 3. This was not possible because many traces did not have lane position of sufficient reliability, and Research Question 3 required a larger sample size than the other research questions.

Data for several positions upstream and along the curve were sampled from the time series data. Models were developed for lane position and speed for both inside (right-hand curve from the perspective of the driver) and outside (left-hand curve from the perspective of the driver), resulting in four models. Lane position was modeled as the offset of the center of the vehicle from the center of the lane. Models were developed using GLS.

Summary of Results

Results indicate that lane position within the curve is influenced by lane position upstream of the curve. The models developed for offset of lane centerline in this study found that drivers who were distracted or glanced away from the roadway tended to shift away from the center of the lane. When driving on the inside of the lane, a driver who was distracted at a particular point within the curve tended to shift 0.14 m to the right by the next point in the curve. When driving on the outside (left-hand curve), a driver who engaged in a non-roadway-related glance at a particular location within the curve was expected to move to the left, or toward the centerline, by 0.13 m at that same point. This confirms the role of distraction in lane keeping.

The models also found that drivers on the inside of a curve tended to move more to the right at the center of curve, while drivers on the outside of a curve were at the furthest point from the centerline at the beginning of the curve. This suggests that drivers may be particularly vulnerable to roadway departures at certain points in the curve negotiation process. Additionally, the lane offset models indicated that age and nighttime are factors in driver lane position.

The model for speeding in the curve found that if drivers are speeding in the upstream, they will also speed in the curve. Drivers of SUVs and pick-up trucks travel on average 2.1 km/h (1.3 mph) faster than drivers of passenger vehicles.

Speeds were predicted to be 0.9 km/h (0.5 mph) lower for each additional 10 years in age for a driver, and drivers who were engaged in a non-roadway-related glance were expected to travel 5.3 km/h (3.3 mph) slower than drivers who were not engage in a non-roadway-related glance. This suggests that drivers whose attention is focused away from the roadway do not maintain longitudinal control.

Implications for Countermeasures

Lane position varies as a function of position within the curve. On the inside of a curve lane, position offset is greatest at the center of the curve. For the outside of the curve lane, position offset is the largest at the beginning of the curve. Additionally, drivers who engaged in eyes-off-roadway distractions tended to shift right of the center of the lane on the inside of the curve.

Both factors indicate that drivers may be more vulnerable to a lane departure at certain points within the curve. As a result, countermeasures such as rumble strips, paved shoulders, and high-friction treatments may ameliorate the consequences of variations in lane position through the curve.

Lane position offset is greater for the outside of the curve during nighttime driving, which suggests that better delineation of curves (edge lines, post-mounted delineators, chevrons) may aid drivers in nighttime curve driving. The models also confirm that drivers who speed upstream are likely to speed within the curve, which suggests that countermeasures that reduce speeds upstream will calm speeds within the curve.

Limitations

The main limitation of this analysis was sample size. Reliable offset data were only available in a subset of the vehicle traces that were reduced. As a result, the number of driver types and roadway features that could be modeled was limited. Consequently, the results are not transferable to all curves or situations. Adding more data to these models may draw out more relationships or strengthen those already found. A more robust data set could also allow for a mixed effects model to be performed, which would allow the findings to be applied toward more curves than those used in the study.

Although the models provided information about factors that result in greater deviation within the lane or higher speeds, the models did not draw correlations between these two factors and increased roadway departure crash risk. It is only assumed that countermeasures that improve lane position or reduce speeds will also reduce roadway departure crashes.

CHAPTER 8

Analysis for Research Question 3

This chapter discusses the third research question: What is the relationship between driver distraction; other driver, roadway, and environmental characteristics; and roadway departure risk?

Research Question 3 investigates how driver behaviors in conjunction with roadway and environmental factors affect the likelihood of a roadway departure on rural two-lane curves. The team had examined a number of different models—including generalized linear models, Bayesian models, and regression tree analysis—during a similar study. That study used NDS data to evaluate roadway departures as part of SHRP 2 Project S01E, Evaluation of Data Needs, Crash Surrogates, and Analysis Methods to Address Lane Departure Research Questions Using Naturalistic Driving Study Data (Hallmark et al. 2011). Logistic regression was determined to be the best model.

Logistic regression models the probability (odds) of a given type of roadway departure based on driver, roadway, and environmental characteristics. *Odds ratios* are the probability that an event happens in relation to the probability that it does not happen. Logistic regression evaluates the association between a binary response and explanatory variables. The natural logarithm of the odds is related to explanatory variables using a linear model. The value of using logistic regression is that the model output (odds ratio) can be easily understood by transportation agencies. As a result, multivariate logistic regression was used to model the probability (odds) of a roadway departure based on driver, roadway, and environmental characteristics. Data are aggregated to the event level for this analysis.

The objective of Research Question 3 was to identify which roadway, environmental, and driver factors are related to roadway departure risk. Models were developed that assessed the probability of a right-side encroachment, probability of a left-side encroachment, probability that the driver will enter the curve 8 km/h (5 mph or more) over the advisory speed if present or posted speed limit if not present, and probability that the driver will enter the curve 16 km/h (10 mph or more) over the advisory speed if present or posted speed limit if not present.

Data Sampling and Modeling Approach for Research Question 3

Data at the event level were used for this analysis. Data were aggregated by epoch: each epoch is one row of data and represents one driver trip through a single curve. The amount of time (decoseconds of data) a driver was traversing the curve was used to normalize epochs with different durations. One row of data (one observation) included information for a section 200 m upstream of the curve and within the curve. Two hundred meters was identified as the curve area of influence in Research Question 1.

A total of 583 observations were included in the analysis. The sample included 57 right-side roadway departures and 40 left-side roadway departures. The advisory speed when present or the posted speed limit when the advisory speed was not present was exceeded by more than 8 km/h (5 mph) for 245 observations and was exceeded by 16 km/h (10 mph) for 123 observations. Data were aggregated to the event level for this analysis.

As mentioned above, the team had previously examined a number of different models—including generalized linear models, Bayesian models, and regression tree analysis—during its work on a similar study (Hallmark et al. 2011). The team considered the objective of Research Question 3 and the type of data being modeled. Additionally, model output was considered because the output of some models is more easily understood than that of others. For instance, logistic regression provides the probability or odds of a certain event happening, which can easily be understood by the stakeholders who are expected to use the results of this research, including state and local transportation agencies. Given these considerations, logistic regression was determined to be the best model.

Variables Used for Research Question 3

Table 8.1 shows the reduced kinematic variables used in this analysis. The two response variables included in the analyses are (1) the probability of a right-side or left-side roadway departure and (2) the probability that the driver entered the curve at 8 km/h or 16 km/h (5 mph or 10 mph) over the curve advisory speed when present or the posted speed limit when not present.

Driver, roadway, and environmental factors were included in the analysis as independent variables. Static roadway characteristics were extracted as described in Chapter 4. Environmental characteristics (e.g., night, raining) were considered to be consistent across the event and were also reduced as described in Chapter 4. Some driver characteristics were also static (e.g., age, gender). While a driver's age changed over the study period, the age of the driver at the time the study commenced was generally used.

Kinematic driver characteristics (e.g., glance location and distraction) were reduced and reported at the same resolution as the time series data from the DAS (10 Hz). Kinematic driver and vehicle characteristics were summarized for 200 m upstream of the point of curvature and through the curve.

Table 8.1. Description of Reduced Kinematic Variables

Variable	Measure
UpOncoming	Fraction of time oncoming vehicles are present for 200 m upstream of the curve (e.g., number of 0.1-s intervals)
Oncoming	Fraction of time oncoming vehicles are present within the curve
UpFollow	Fraction of time oncoming subject vehicles following lead vehicle 200 m upstream of the curve
Follow	Fraction of time oncoming subject vehicles following lead vehicle
Up_Spd	Speed averaged over 200 m upstream of the curve (m/s)
Up_Std	Standard deviation of speed for 200 m upstream of the curve
Speed	Speed averaged over curve (m/s)
Std	Standard deviation of speed within the curve
Ent_Spd	Speed at which vehicle entered curve (m/s)
UP_FR	Fraction of time driver glance location is forward roadway for 200 m upstream of the curve
FR	Fraction of time driver glance location is forward roadway over the curve
UP_SA	Fraction of time driver glance location is to roadway-related tasks for 200 m upstream of the curve
SA	Fraction of time driver glance location is to roadway-related tasks over the curve
UP_NR	Fraction of time driver glance location is to non-roadway-related tasks for 200 m upstream of the curve
NR	Fraction of time driver glance location is to non-roadway-related tasks over the curve
UP_PASS	Fraction of time driver glance location is away from roadway on passenger-related tasks for 200 m upstream of the curve
PASS	Fraction of time driver glance location is away from roadway on passenger-related tasks over the curve
UP_InVeh	Fraction of time driver glance location is away from roadway on in-vehicle-control-related tasks for 200 m upstream of the curve
InVeh	Fraction of time driver glance location is away from roadway on in-vehicle-control-related tasks over the curve
UP_CELL	Fraction of time driver glance location is away from roadway on cell phone–related tasks for 200 m upstream of the curve
CELL	Fraction of time driver glance location is away from roadway on cell phone–related tasks over the curve
UP_PerHY	Fraction of time driver glance location is away from roadway on personal hygiene–related tasks for 200 m upstream of the curve
PerHY	Fraction of time driver glance location is away from roadway on personal hygiene–related tasks over the curve
UP_EAT	Fraction of time driver glance location is away from roadway on eating/drinking-related tasks for 200 m upstream of the curve
EAT	Fraction of time driver glance location is away from roadway on eating/drinking-related tasks over the curve
Up_AVG_SR	Average length of glance away from forward roadway to roadway-related locations (seconds) for 200 m upstream of curve
AVG_SR	Average length of glance away from forward roadway to roadway-related locations (seconds) through curve
Up_AVG_NR	Average length of glance away from forward roadway (seconds) to non-roadway-related locations for 200 m upstream of curve
AVG_NR	Average length of glance away from forward roadway (seconds) to non-roadway-related locations through curve

These characteristics were reduced to the event level (see Table 8.1). For instance, vehicle speed for each 0.1-s interval over the curve was averaged.

Random effect variables were included to account for multiple samples for the same driver or same curve. Separate models were developed for right- and left-side roadway departures because they are likely to be affected by different factors and are two distinct events.

Description of Analytical Approach and Results for Research Question 3

Right-Side Encroachment

Logistic regression was used to model the odds of a right-side encroachment (0 for no right-side encroachment, 1 for right-side encroachment) for each event indexed by i as a random variable, γ_i, which follows a Bernoulli distribution with probability of departure, p_i.

$$\gamma_i \sim Bernoulli(p_i)$$

For the logistic regression, the log odds of a right-side encroachment were modeled as follows:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \alpha_i + \gamma_i$$

$$\alpha_i \sim Normal(0, \sigma_d^2)$$

$$\gamma_i \sim Normal(0, \sigma_c^2)$$

where

x_1 = proportion of time the driver is glancing at the forward roadway 200 m upstream of the curve;
x_2 = indicator for curve direction (0 = outside; 1 = inside);
x_3 = radius of the curve (m);
x_4 = indicator for presence of a guardrail (0 = not present; 1 = present);
x_5 = indicator for presence of a curve warning sign, either a W1-6 sign or curve advisory sign (0 = not present; 1 = present);
α_i = random effect for subject; and
γ_i = random effect for curve.

This model was chosen as the best fit model for the data by model selection using PROC LOGISTIC in SAS and by comparing the AIC/BIC values to various other models that were examined.

All models were fit using the glmer() command in the lme4 package in R. The fitted model parameters, p-values, and 90% Wald confidence intervals are shown in Table 8.2.

Table 8.2. Parameter Estimates for Right-Side Encroachments

Parameter	Estimate	*p*-value	5%	95%
β_0	–1.4039	0.1998	–3.2049	0.3972
β_1	–1.9968	0.0388	–3.5866	–0.407
β_2	1.9116	2.57E-06	1.2194	2.6039
β_3	–4.00E-04	0.0399	–8.00E-04	–1.00E-04
β_4	–1.121	0.1072	–2.2656	0.0237
β_5	0.2337	0.6144	–0.5293	0.9967
σ_d^2	1.1754	NA	NA	NA
σ_c^2	0.4243	NA	NA	NA

Wald intervals were calculated as follows:

$$\hat{\beta}_i \pm z_{.95} * s_i$$

where

$\hat{\beta}_i$ = the estimate of the parameter given above;
$z_{.95}$ = the 95th percentile of the standard normal distribution; and
s_i = the standard error of the estimate (not given).

The odds ratios for this model, which are equivalent to $\exp(\beta_i)$ for $i = 1, \ldots, 6$, are given in Table 8.3. For the dummy variables, the odds of a right-side encroachment change by a factor of $\exp(\beta_i)$ when the object (traveling on the inside of the curve, etc.) is present relative to when it is not present. For the numeric variables, the odds of a right-side encroachment change by a factor of $\exp(\beta_i)$ when the covariate x_i increases by 1 unit. Table 8.3 shows the 90% Wald intervals for these estimates, which are calculated by exponentiating the boundaries of the confidence intervals in Table 8.2.

Because Up_FR is a proportion, a more appropriate estimate of the odds ratio might correspond to the change in odds of lane departure when forward glance time increases by 0.10, or

Table 8.3. Confidence Intervals for Right-Side Encroachments

Variable	Odds Ratio Estimate	5%	95%
Up_FR	0.1358	0.0277	0.6656
Direction (inside versus outside)	6.7642	3.385	13.5167
Radius	0.9996	0.9992	0.9999
Guardrail (present versus not present)	0.326	0.1038	1.024
Curve warning (present versus not present)	1.2632	0.589	2.7093

10% of total time upstream of the curve. This value is given by the following:

$$\exp(0.10\beta_1) = 0.8190$$

Thus, the odds of lane departure increase by a factor of $1/0.8190 = 1.221$ when the proportion of forward glance time upstream of the curve decreases by 0.10.

The results also indicate that a right-side lane departure is 6.8 times more likely on the inside of a curve compared with the outside of the curve.

Lane departures are slightly more likely (1.3 times) for curves with any type of curve advisory sign (including W1-6). It is unlikely that the presence of the warning sign leads to increased probability of a right-side encroachment. Rather, it is likely that advisory signs are more likely to be present on curves of a certain type (i.e., those with sight distance issues, sharper curves), and encroachments are also more likely for those road types. Additionally, the results suggest that the simple presence of curve warning signs does not mitigate roadway departures.

A statistically significant but small correlation exists between radius of curve and probability of a right-side encroachment.

Drivers were 0.33 times less likely to have a right-side encroachment on roadways where a guardrail is present. A guardrail is used to decrease the severity of a crash when a vehicle leaves the roadway. It is not a countermeasure to prevent roadway departures. The presence of a guardrail may suggest to the driver that roadway conditions are less safe, resulting in better driver attention. Additionally, few delineation countermeasures (e.g., chevrons) were present in the curves included in the analysis. As a result, a guardrail may provide some delineation of the curve, which provides feedback to the driver about the sharpness of the curve.

Left-Side Encroachment

The log odds of left-side encroachment were modeled as follows:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \alpha_i + \gamma_i$$

$$\alpha_i \sim Normal\left(0, \sigma_d^2\right)$$

$$\gamma_i \sim Normal\left(0, \sigma_c^2\right)$$

where

x_1 = dummy variable for driver gender (0 = female; 1 = male);
x_2 = dummy variable for the direction of the curve (0 = outside; 1 = inside);
x_3 = radius of the curve;
α_i = random effect for drivers; and
γ_i = random effect for curve.

Table 8.4. Parameter Estimates for Left-Side Encroachments

Parameter	Estimate	*p*-value	5%	95%
β_0	–3.2435	2.00E-04	–4.6796	–1.8074
β_1	1.4888	0.0753	0.112	2.8656
β_2	–2.2653	7.00E-04	–3.3661	–1.1645
β_3	–7.00E-04	0.06	–0.0012	–1.00E-04
σ_d^2	1.879	NA	NA	NA
σ_c^2	2.383	NA	NA	NA

Parameter estimates, *p*-values, and 90% Wald confidence intervals are shown in Table 8.4.

Odds ratios and 90% Wald confidence intervals are shown in Table 8.5.

As noted, males are more than four times more likely to have a left-side encroachment, and drivers traveling on the inside of the curve are 0.1 times less likely to have a left-side encroachment than drivers traveling on the outside of the curve. The impact of radius was statistically significant but minor.

Left-side encroachments are likely to be drivers who "cut the curve." Although driver intent is difficult to determine, in several cases the driving manner as evidenced in the forward videos suggested that the driver was intentionally crossing the centerline.

Probability of Exceeding Posted or Advisory Speed by 8 km/h (5 mph)

The log odds of a vehicle traveling 8 km/h (5 mph) or more over the posted or advisory speed limit was also modeled using logistic regression as follows:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \alpha_i + \gamma_i$$

$$\alpha_i \sim Normal\left(0, \sigma_d^2\right)$$

$$\gamma_i \sim Normal\left(0, \sigma_c^2\right)$$

Table 8.5. Confidence Intervals for Left-Side Encroachments

Variable	Odds Ratio Estimate	5%	95%
Gender (male versus female)	4.4318	1.1186	17.5587
Direction (inside versus outside)	0.1038	0.0345	0.3121
Radius	0.9993	0.9988	0.9999

where

x_1 = driver age (years);
x_2 = fraction of time following another vehicle 200 m upstream of the curve;
x_3 = average speed 200 m upstream of the curve (m/s);
x_4 = indicator for roadway markings (0 = visible markings; 1 = obscured markings/not present);
x_5 = indicator for visibility (0 = clear; 1 = any reduced visibility);
x_6 = the radius of the curve (m);
x_7 = indicator for presence of a paved shoulder (0 = not present; 1 = present);
x_8 = indicator for raised pavement markings (0 = not present; 1 = present);
α_i = random effect for drivers; and
γ_i = random effect for curve.

Again, the categorical variables can be thought of as having more than one coefficient, with the interpretations of estimates and odds ratios the same as in the previous section.

Parameter estimates, *p*-values, and 90% Wald confidence intervals are shown in Table 8.6.

Odds ratios and 90% Wald confidence intervals are shown in Table 8.7.

Results show that drivers who are following other vehicles or driving under reduced visibility conditions are less likely to enter the curve at 8 km/h (5 mph) or more over the posted or advisory speed. Drivers traveling at higher speeds upstream are much more likely to enter the curve at 5 mph over the speed limit, as expected. Additionally, when pavement markings are obscured or not present, drivers are significantly more likely to enter the curve more than 8.0 (5 mph) over the posted/advisory speed. Lane line markings may provide curve delineation, which aids drivers in gauging the sharpness of the curve so that they are better able to select curve entry speeds.

Table 8.6. Parameter Estimates for 5 mph Over the Speed Limit

Parameter	Estimate	5%	95%	*p*-value
β_0	–26.0071	–34.617	–17.3973	6.75E-07
β_1	–0.0683	–0.1219	–0.0146	0.0365
β_2	–2.0832	–3.2748	–0.8916	0.004
β_3	1.8159	1.4302	2.2016	9.64E-15
β_4	7.2735	3.8652	10.6818	4.00E-04
β_5	–4.94	–7.4111	–2.4688	0.001
β_6	–0.0015	–0.0022	–8.00E-04	3.00E-04
β_7	–10.3751	–15.0862	–5.664	3.00E-04
β_8	–7.3919	–10.3264	–4.4574	3.42E-05
σ_d^2	41.215	NA	NA	NA
σ_c^2	4.831	NA	NA	NA

Table 8.7. Confidence Intervals for 8 km/h (5 mph) Over the Speed Limit

Parameter	Estimate	5%	95%
Age	0.934	0.8852	0.9855
UpFollow	0.1245	0.0378	0.41
UpSpd	6.1465	4.1795	9.0393
Markings	1441.5916	47.7128	43556.112
Visibility	0.0072	6.00E-04	0.0847
Radius	0.9985	0.9978	0.9992
PvdShd	3.11998E-05	2.80638E-07	0.0035
RPM	6.00E-04	3.27568E-05	0.0116

The odds ratio estimates for the paved shoulder and raised paved markings indicator variables are extremely small, though they are still significantly different from zero in the model. This suggests that drivers are less likely to speed when paved shoulders or raised pavement markings are present. However, these variables are best used within the whole model, instead of being considered separately, as their estimates are almost nonsensical. However, their extreme values do indicate that their presence significantly decreases the odds of speeding more than 5 mph over the posted/advisory speed.

Probability of Exceeding Posted or Advisory Speed by 16 km/h (10 mph)

The log odds of a driver exceeding the posted or advisory speed by 16 km/h (10 mph) or more were modeled using the following equation:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \alpha_i + \gamma_i$$

$$\alpha_i \sim Normal\left(0, \sigma_d^2\right)$$

$$\gamma_i \sim Normal\left(0, \sigma_c^2\right)$$

where

x_1 = average speed 200 m upstream of the curve (m/s);
x_2 = average amount of time driver glance is away from the road and engaged in driving tasks;
x_3 = radius of the curve (m);
x_4 = indicator for paved shoulder (0 = unpaved; 1 = paved);
x_5 = indicator for raised pavement markings (0 = not present; 1 = present);
α_i = random effect for drivers; and
γ_i = random effect for curve.

Table 8.8. Parameter Estimates for 16 km/h (10 mph) Over the Speed Limit

Parameter	Estimate	*p*-value	5%	95%
β_0	–57.7719	6.77E-07	–76.9007	–38.6432
β_1	3.3555	1.70E-10	2.4912	4.2198
β_2	–1.8275	0.0485	–3.3513	–0.3038
β_3	–0.0028	0.0429	–0.005	–5.00E-04
β_4	–24.6413	0.0048	–39.002	–10.2807
β_5	–16.3668	0.0259	–28.4486	–4.285
σ_d^2	339.603	NA	NA	NA
σ_c^2	2.904	NA	NA	NA

Parameter estimates, *p*-values, and 90% Wald confidence intervals are provided in Table 8.8.

The odds ratios and 90% Wald confidence intervals are shown in Table 8.9.

As noted in Table 8.9, the average speed upstream significantly increases the likelihood that a driver will also exceed the curve advisory/posted speed limit by 16 km/h (10 mph) or more. The length of the glance away from the forward roadway to roadway-related tasks (e.g., glances at rearview mirror) decreases the likelihood that drivers will exceed the posted/advisory speed. This indicates that drivers may tend to slow down when they engage in longer glances away from the forward roadway.

The probability of exceeding the posted/advisory speed is correlated to radius. As radius increases, drivers are less likely to exceed the posted/advisory speed. Curves with smaller radii are more likely to have an advisory speed, and presumably drivers are more likely to exceed lower speed limits.

The odds ratio estimates for the paved shoulder and raised paved markings indicator variables are extremely small, though they are still significantly different from zero in the model. This suggests that drivers are less likely to speed when paved shoulders or raised pavement markings are present. However, these variables are best used within the whole model, instead of being considered separately, as their estimates are almost nonsensical. However, their extreme values do indicate that their presence significantly decreases the odds of speeding over 16 km/h (10 mph).

Table 8.9. Confidence Intervals for 16 km/h (10 mph) Over the Speed Limit

Variable	Odds Ratio Estimate	5%	95%
UpSpd	12.0763	12.0763	68.0172
Avg_SA	0.035	0.035	0.738
Radius	0.995	0.995	0.9995
PvdShd	1.15252E-17	1.15252E-17	3.42885E-05
RPM	4.41499E-13	4.41499E-13	0.0138

Summary of Crash/Near-Crash Events

Only one crash and three near crashes were present in the data set. Therefore, a model could not be developed for this research. All three near crashes appeared to be near rear-end collisions, with the roadway departure caused by the driver swerving to avoid the potential rear-end crash.

A review of the data for the crash indicated that speeding was likely a major factor because the driver was 15 km/h (9 mph) over the posted speed limit of 72 km/h (45 mph) and 55 km/h (34 mph) over the curve advisory speed of 32 km/h (20 mph). The driver was not engaged in any distracting tasks and only glanced away from the forward view once (a glance to the steering wheel, which lasted 0.6 s). The crash occurred between midnight and 3:00 a.m. There was some evidence of drowsiness because the driver was resting his head on his right arm or hand.

The radius of curve was 50 m, and no shoulders, chevrons, or other countermeasures were present.

Summary and Discussion

The objective of Research Question 3 was to assess the relationship between driver, roadway, and environmental factors and risk of a roadway departure. The crash surrogates used for this research question were probability of a right-side encroachment, probability of a left-side encroachment, probability that the driver exceeded the posted or advisory speed by 5 mph or more, and probability that the driver exceeded the posted or advisory speed by 10 mph or more. Logistic regression was used to model observations at the event level.

Key Findings

Four different models were developed. The model for right-side encroachments indicated that the probability of a right-side encroachment increases as drivers spend less time glancing at the forward roadway. The results also indicate that a right-side lane departure is 6.8 times more likely on the inside of a curve compared with the outside of the curve. Lane departures are slightly more likely (1.3 times) for curves with any type of curve advisory sign (including W1-6). A statistically significant but small correlation exists between radius of curve and probability of a right-side encroachment. Drivers were 0.33 times less likely to have a right-side encroachment on roadways with a guardrail.

The model for left-side encroachments indicated that males are more than four times more likely to have a left-side lane departure, and drivers traveling on the inside of the curve are 0.1 times less likely to have a left-side encroachment than drivers traveling on the outside of the curve. The impact of radius was statistically significant but minor.

The probability that a driver will be 8 km/h (5 mph) or more over the posted/advisory speed is higher for younger drivers, when drivers have a higher average speed upstream, and when edge line markings are obscured or not present. The amount of time a driver spends following another vehicle, presence of lower visibility conditions, and presence of paved shoulders and RPMs decreases the probability that a driver will enter the curve 8 km/h (5 mph) or more over the posted/advisory speed.

The probability that a driver will be 16 km/h (10 mph) or more over the posted/advisory speed is higher when drivers have a higher average speed upstream. The probability is lower when the average glance at roadway-related tasks is longer and when paved shoulders and RPMs are present.

Implications for Countermeasures

The presence of warning signs increased the likelihood of a right-side encroachment. It is unlikely that the presence of a warning sign itself increases the probability. Rather, it is likely that advisory signs are more likely to be present on curves of a certain type (i.e., those with sight distance issues, sharper curves), and encroachments are also more likely for those road types. Additionally, the results suggest that the simple presence of curve warning signs does not mitigate roadway departures.

The presence of a guardrail decreased the probability of a right-side encroachment. The purpose of a guardrail is to mitigate the consequences of a driver leaving the roadway rather than to keep the driver from leaving the roadway. Consequently, a guardrail in and of itself does not mitigate roadway departures. The presence of a guardrail may suggest to the driver that roadway conditions are less safe, resulting in better driver attention. Additionally, few delineation countermeasures (e.g., chevrons) were present in the curves included in the analysis. As a result, a guardrail may provide some delineation of the curve, which provides feedback to the driver about the sharpness of the curve.

The probability that a driver would exceed the posted/advisory speed by 5 mph or more was higher for curves with obscured/missing edge lines. Presence of RPMs decreased the probability of exceeding the posted/advisory speed by 8 km/h and 16 km/h (5 mph and 10 mph). Taken together, these results indicate that better curve delineation may allow drivers to better gauge upcoming changes in roadway geometry, resulting in better speed selection and decreased risk of a roadway departure, and it may help decrease speed. Delineation countermeasures include chevrons, the addition of reflective panels to existing chevron posts, reflective barrier delineation, RPMs, post-mounted delineators, edge lines, and wider edge lines.

The speed models suggest that driver age and upstream speed have a significant impact on drivers' speed within a curve. As a result, speed management countermeasures that affect tangent speed will also decrease curve speeds. The results also indicate that speed management is appropriate to get drivers' attention before entering a curve. Countermeasures specifically targeted to reduce speed on curves include dynamic speed feedback signs, on-pavement curve warning signs, and flashing beacons.

Limitations

The most significant limitations are sample size and representation of different curve and driver characteristics. More than 700 potential curves were initially identified. This represented a wide range of roadway characteristics and countermeasures. However, some countermeasures, such chevrons and rumble strips, were not widely available in the study areas, and some countermeasures, such as post-mounted delineators, were not available at all. Additionally, only one-third of the full NDS data set was available for query at the time the data request was made, and data were only found for 148 curves, which reduced the number of roadway characteristics that could be included. A total of 583 observations were included in the analysis. However, only 57 right-side roadway departures and 40 left-side roadway departures were present.

Another limitation is that crashes/near crashes were not available, so the relationship between encroachments or speed and roadway departure crash risk could not be established.

Additionally, although a model was developed for left-side encroachments, this model is likely to include drivers who cut the curve. Although it is difficult to determine driver intent, in several cases the driving manner as evidenced in the forward videos suggested that the driver was intentionally crossing the centerline.

CHAPTER 9

Analysis for Research Question 4

This chapter discusses the fourth research question: Can lane position at a particular state be predicted as a function of position in a prior state?

Research Question 4 focuses more specifically on driver response to changing roadway characteristics and traffic conditions. Time series models were developed to incorporate the dynamic process of information acquisition and response as a driver negotiates a curve. The analysis evaluated the influence of roadway geometries or traffic conditions on drivers' lane-keeping behavior. For example, drivers on a rural two-lane roadway tend to have larger lane deviation from the centerline when there is an oncoming vehicle.

Two types of dynamic linear models (DLMs) were built in this study to describe and explain the curve negotiation process: DLM with intervention analysis and DLM with autoregression and moving average (ARMA). The DLM with intervention analysis was mainly used for explanatory purposes, which related lane offset to curve characteristics and traffic conditions. The DLM with ARMA was mainly used for forecasting purposes, which could be used for roadway departure warning systems.

It should be noted that Research Question 4 is similar to Research Question 2. Research Question 4 uses a different statistical method to explore whether driver/vehicle characteristics in one time state during curve negotiation can be modeled from previous states using a time series analysis. Research Question 4 is included as a separate research question to simplify discussion of methodology and results.

The objective of Research Question 4 was to determine the feasibility of the approach because there was not sufficient time to conduct a full-scale analysis.

Description of Analytical Approach for Research Question 4

A DLM was used to analyze driver behavior. The DLM can account for the autocorrelation of the observations in time. It is a flexible model that allows the inclusion of explanatory variables and stochastic time components in the same model. The explanatory variables can evolve over time. The popular Box-Jenkins autoregressive integrated moving average (ARIMA) model was not used in this study, because the underlying process of the Box-Jenkins model is assumed not to change over time; driving behavior varies in different roadway segments with different speed limits, roadway geometries, and weather conditions. The DLM is a better model in this case because the model can consistently update the model parameters based on the modeling errors in previous steps. In other words, the model can evolve over time based on its past observations and can be adapted to different situations. The general form of DLM can be written as follows:

Observation equation $Y_t = F_t\theta_t + \nu_t$, with $\nu_t \sim N_1(0, V_t)$

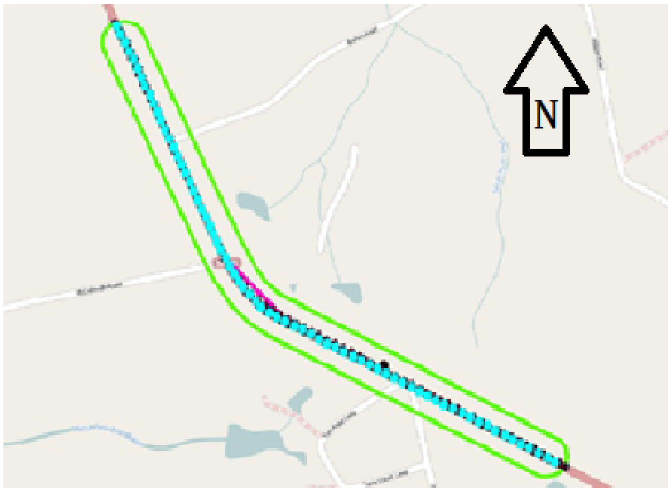
State evolution equation $\theta_t = G_t\theta_{t-1} + \omega_t$, with $\omega_t \sim N_p(0, W_t)$

Initial prior $(\theta_0 \mid D_0) \sim N(m_0, C_0)$, where (m_0, C_0) is fixed and $D_t = \{Y_t, D_{t-1}\}$

The model assumes that the underlying state, θ_t, evolves smoothly over time as an autoregressive process and that the observation at time, t, is a smooth function of the state. The state evolution equation, as formulated above, is a function of an underlying process that is unobserved. Explanatory variables can be included as part of the underlying process driving the observation equation by including a linear combination of the explanatory variables in the state evolution equation. Coefficients F_t and G_t are often assumed to be constant over time, but they can also be time dependent.

Data Sampling and Segmentation Approach for Research Question 4

One curve in North Carolina was selected that has lane position, speed, and pedal position data that were determined to be sufficiently reliable for the model. The selected trip is



Source: World_Street_Map (Esri, DeLorme, NAVTEQ, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China [Hong Kong], Esri [Thailand], and TomTom, 2013).

Figure 9.1. Sample trip highlighted in ArcGIS.

shown in Figure 9.1. The sample curve is a single curve with a radius of 1,128 m. The driver was driving from northwest to southeast on the outside lane on a rural two-lane highway. The speed limit is 55 mph. The trip occurred at night, and there was only one oncoming vehicle in this trip.

Variables Used for Research Question 4

Time series data output by the DAS with additional variables added were used for this model. The distance from the left wheel to the left lane marking was used as a dependent variable. A positive value for the left distance indicates that the left wheel is within the lane; a negative value for the left distance indicates that the left wheel crossed the centerline marking. The raw data were collected at every 0.1 s but were aggregated to the 1-s level for this analysis. The time series plot of the raw data for the left distance is shown in Figure 9.2. The plot also labels the area of influence under the oncoming vehicle and the curve. The figure shows that the vehicle had a larger left distance, so it moved away from the centerline when there was an oncoming vehicle. In contrast, the vehicle moved closer to the centerline when the driver was driving inside the curve.

Results for Research Question 4

This analysis focused on the use of DLM for intervention analysis and forecasting. The next section introduces the intervention analysis. The subsequent section focuses on the use of DLM for forecasting.

Intervention Analysis with Dynamic Linear Model

The objective of DLM is to describe and explain the lane position of the vehicle in the curve negotiating process. The proposed model assumes an additive model in which the lateral position is the sum of normal driving positions, the lane deviation due to the oncoming vehicle, and the lane deviation due to the curve. The influence of the oncoming vehicle and the curve are introduced into this model as intervention variables $W_{v,t}$ and $W_{c,t}$, respectively, as follows:

$$y_t = \mu_t + \beta_{v,t} W_{v,t} + \beta_t W_{c,t} + \varepsilon_t \qquad \varepsilon_t \sim \mathrm{NID}\left(0, \sigma_\varepsilon^2\right)$$

$$\mu_{t+1} = \mu_t + \xi_t \qquad \xi_t \sim \mathrm{NID}\left(0, \sigma_\xi^2\right)$$

$$\beta_{v,t+1} = \beta_{v,t} + \tau_t \qquad \tau_t \sim \mathrm{NID}\left(0, \sigma_\tau^2\right)$$

$$\beta_{c,t+1} = \beta_{c,t} + \rho_t \qquad \tau_t \sim \mathrm{NID}\left(0, \sigma_\rho^2\right)$$

For $t = 1 \ldots n$, where μ_t is a stochastic-level variable at time t, the model represents the normal driving positions without the influence of the oncoming vehicle and the curve. $W_{v,t}$ is a dummy variable with 0 for the absence and 1 for the existence of an oncoming vehicle. $W_{c,t}$ is a dummy variable with 0 for outside the curve and 1 for inside the curve. The coefficients $\beta_{v,t}$ and $\beta_{c,t}$ are the coefficient matrices of these intervention effects. The variable ε_t is the random noise in the observation equation. The model parameters ξ_t, τ_t, and ρ_t are the random noise in the state equations.

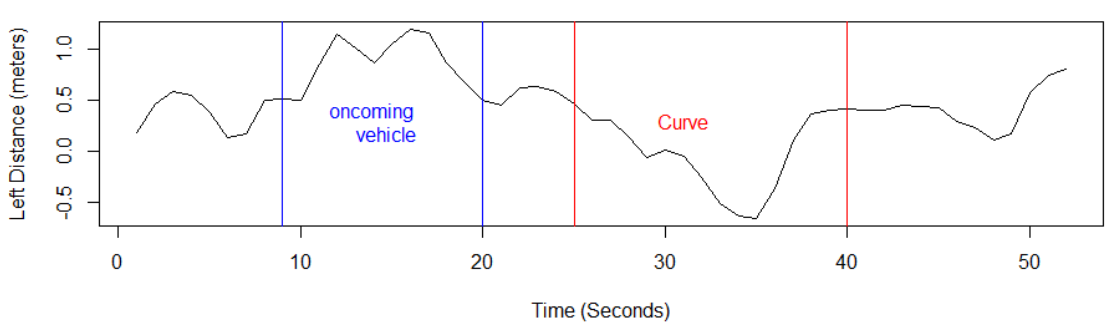


Figure 9.2. Time series plot of distance of left wheel to left lane marking.

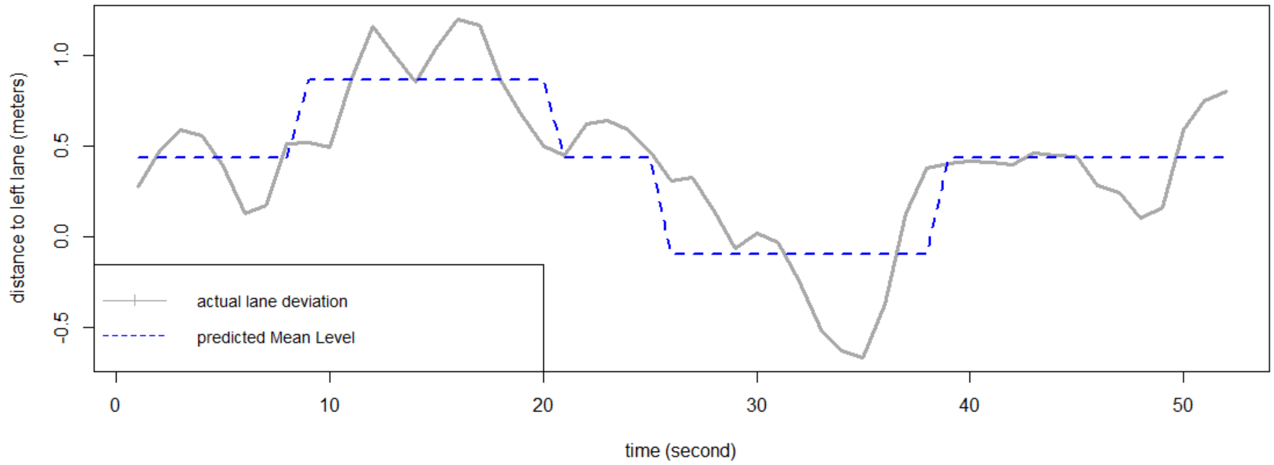


Figure 9.3. Actual lane deviation versus predicted lane deviation based on DLM.

The model was estimated using maximum likelihood estimation based on the DLM package in R. The predicted mean level of the model is plotted in Figure 9.3. This model assumes the variance for the state equations (ξ_t, τ_t, ρ_t) to be zero. In this way, the model is forced to fit a straight line for the mean effect. The influence of the oncoming vehicle is treated as an intervention effect and causes a level shift to a larger left distance, whereas the curve causes a level shift to a smaller left distance, as illustrated in Figure 9.3.

After the model was fitted with the DLM, the model was further decomposed into three components: the mean level at normal driving, the intervention effect due to an oncoming vehicle, and the intervention effect due to the curve. The decomposition of each effect allows the evaluation of the three components separately, as shown in Figure 9.4.

The top panel in Figure 9.4 illustrates the mean level of lane deviation as if there were no intervention effect from the oncoming vehicle and the curve. It also represents the mean lane position for normal driving conditions, in which the left edge of the vehicle is approximately 0.44 m (1.44 ft) from the centerline (to the right of the centerline). The middle panel shows the influence of an oncoming vehicle on the lane deviation. The positive sign means the vehicle was moving away from the centerline for an additional 0.43 m (1.41 ft) (a total

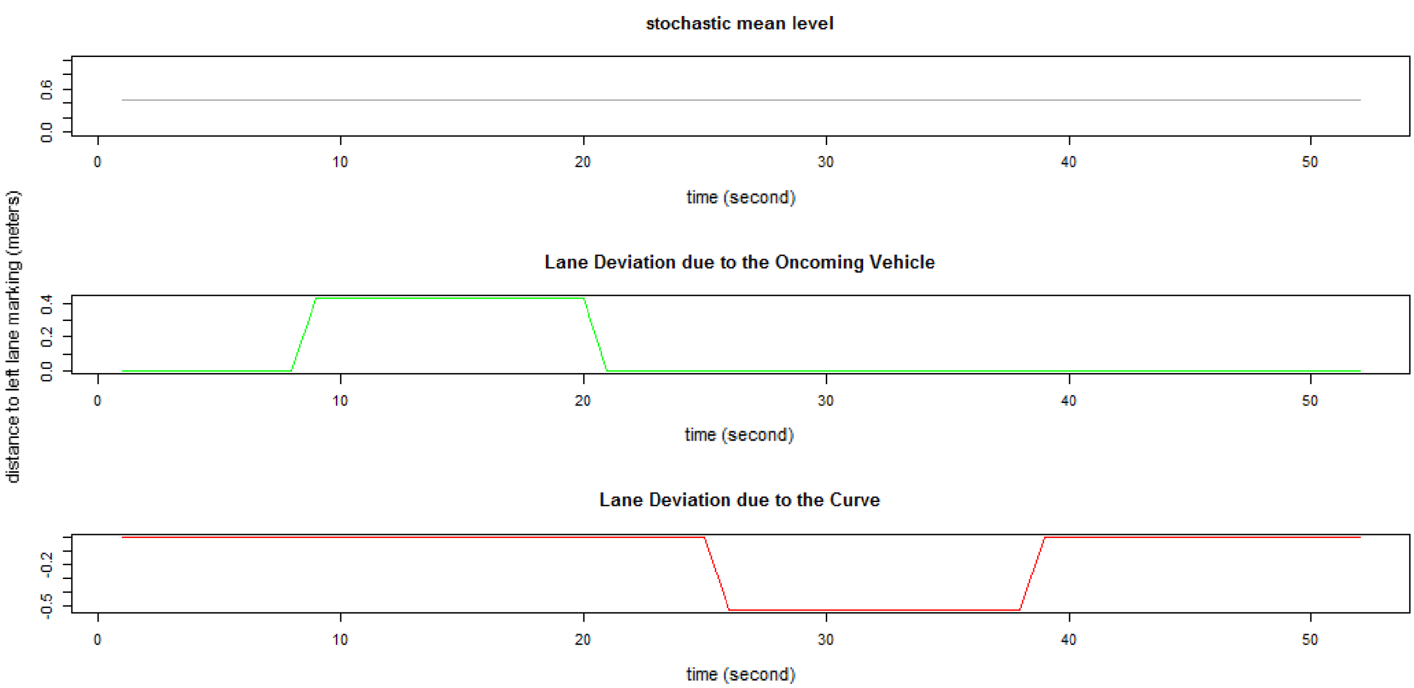


Figure 9.4. Decomposition of lane deviation data into three components: Stochastic mean effect, lane deviation due to oncoming vehicle, and lane deviation due to curve.

of 0.77 m to the right of the centerline) to avoid the conflict with the oncoming vehicle. The bottom panel illustrates the intervention effects of the curve on the lane position. The effect of the curve on the lane position is –0.53 m (1.73 ft) to the left of the normal driving position, which places the vehicle about 0.09 m (0.30 ft) beyond the centerline (i.e., the left edge of the vehicle crosses the centerline by 0.1 m).

Forecasting with Dynamic Linear Model

The second model is a DLM representation of ARMA model. The classical ARMA (p, q) process model can be defined as the following formula:

$$Y_t = \sum_{j=1}^{p} \phi_j Y_{t-j} + \sum_{j=1}^{q} \psi_j \varepsilon_{t-j} + \varepsilon_t$$

where

Y_{t-j} = the past observation at time $t-j$;

ε_{t-j} = the error of the model at time $t-j$; and

ε_t = the residual of the model at time t.

This model predicts the future position of the vehicle based on the past p number of observations and q number of modeling errors. The best model for this time series data is the ARMA (2, 1) model, as follows:

$$Y_t = 0.89Y_{t-1} + 0.06Y_{t-2} + 0.98\varepsilon_{t-1} + \varepsilon_t$$

The variance of the model residuals is 0.015. The one-step-ahead prediction with 95% confidence interval is plotted in Figure 9.5. The figure shows that 95% of the one-step-ahead predictions will fall into this interval. The width of the confidence interval is approximately 0.4 m. The plot shows that the model predicted the future observations accurately.

The model fits and residuals were checked with a variety of tests. The Q-Q plot and Shapiro-Wilk normality test were used to check the normality of the residuals. The Q-Q plot followed a straight line, and the p-value for the Shapiro-Wilk normality test is 0.56, which is higher than the 0.05 confidence level. Therefore, the model satisfies the normality assumption.

The standardized residuals are plotted in Figure 9.6. Most of the standardized residuals in the plot are lower than 3, which means no significant outlier was detected. Both an autocovariance function (ACF) and Ljung-Box statistics were used to check the correlations in the residuals. The ACF values for all lags are within the limits. The p-values of the Ljung-Box statistics are all above 0.05, which indicates no significant autocorrelation in the residuals. The forecast of the model is also reasonably similar to the actual values. Therefore, it was concluded that the DLM fit the data well. Again, the limitation is that the model can only be used to treat one time series data row at a time. The model is also not appropriate to generalize this result to other drivers on other curves.

Summary and Implications

This chapter has described a time series analysis of driver behavior on rural two-lane curves using SHRP 2 NDS data. DLM was used to analyze driver behavior as a dynamic process. Distance from the centerline was chosen as a surrogate to represent roadway departure risks. Two types of DLM were used to analyze the time series data: DLM with intervention analysis and DLM with ARMA for forecasting.

The intervention analysis evaluated the influence of the oncoming vehicle and the curve on lane position. The two effects were included in the DLM as intervention effects. The model was decomposed into three components: the mean level of lane position, the lane deviation due to the oncoming vehicle, and the lane deviation due to the curve.

This analysis found that the average distance from the centerline to the left edge of the vehicle under normal driving conditions is approximately 0.44 m (1.44 ft) to the right of the centerline. The vehicle moved away from the centerline by an additional 0.43 m (1.41 ft) when there was an oncoming

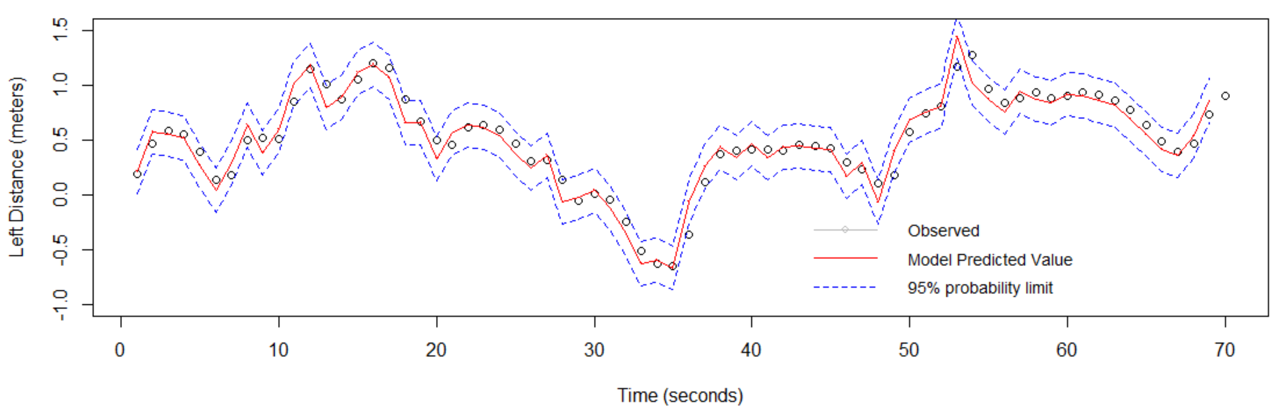


Figure 9.5. Predicted left distance with 95% confidence interval.

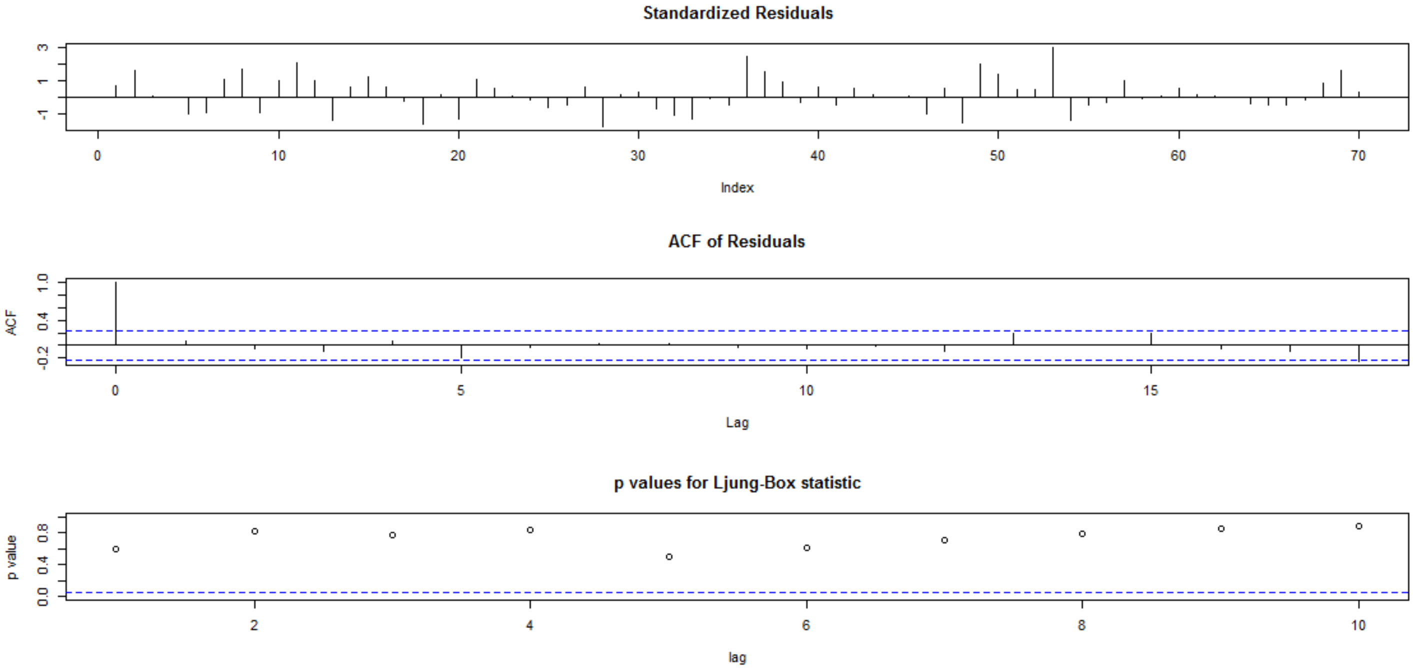


Figure 9.6. Diagnostics of residuals: Standardized residuals, autocovariance function of residuals, and p-values for Ljung-Box statistics.

vehicle (a total of 0.77 m to the right of the centerline). At the inside of the curve, the vehicle moved closer to the centerline by about 0.53 m (1.73 ft), which placed the left vehicle edge across the centerline by 0.09 m (0.30 ft).

However, these findings are based on one sample trip only and cannot be generalized to other drivers and curves. The overall safety benefit of this finding is arguable. On the one hand, larger lane deviation from the centerline increased the probability of a roadway departure crash. On the other hand, it decreased the likelihood of a head-on collision with an oncoming vehicle. Therefore, the overall safety benefits of the lane position should be further evaluated and discussed in a future study.

The second model, DLM with ARMA, successfully fitted the sample trip for forecasting purposes. The diagnostics of the model residuals indicated that the fitting is adequate for the time series data. The model predicted the future position of the vehicle based on the past observations and errors. One-step-ahead prediction showed that the predicted values are very close to the actual observations. The 95% confidence interval was also plotted with the forecast values. Therefore, the research team concludes that the model fit the data well and could potentially be used for roadway departure crash warning systems.

The model results are only applicable to the scenarios tested and cannot be extrapolated to other drivers or other curves. The intent of Research Question 4 was to show proof of concept. However, results suggest that the time series data can be used to model driver lateral control as a function of external variables (oncoming vehicles, position within the curve). This indicates that lane position data from the SHRP 2 NDS that are sufficiently reliable could be used in development of collision warning system algorithms.

It should be noted, however, that the models predicted lane position at a high resolution (i.e., 0.09 m). The accuracy of offset and other lane position variables has not yet been published. The accuracy also depends on the quality of the forward view, quality of lane lines, and other factors. As a result, 0.1 m (0.3 ft) is likely smaller than the accuracy of the variables used to calculate position from the lane edge. Therefore, whether a vehicle crosses a lane line and the magnitude of the crossing should be evaluated cautiously. The results do show that the model can be used to predict position and shifting position due to external variables.

CHAPTER 10

Summary and Recommendations

Summary

Over half of motor-vehicle fatalities are roadway departures. Rural horizontal curves are of particular interest because they have been correlated with overall increased crash occurrence. Although transportation agencies expend significant resources to address the problem, the interaction between the driver and roadway environment is not well understood. As a result, it is difficult to select appropriate countermeasures.

To address this knowledge gap, data from the SHRP 2 NDS and RID were used to develop relationships between driver, roadway, and environmental characteristics and risk of a road departure on rural curves. Only curves on rural two-lane paved roadways with posted speed limits of 64 km/h to 97 km/h (40 mph to 60 mph) were included.

The research was tailored to address four fundamental research questions:

1. What defines the curve area of influence?
2. What defines normal behavior on curves?
3. What is the relationship between driver distractions; other driver, roadway, and environmental characteristics; and risk of roadway departure?
4. Can lane position at a particular state be predicted as a function of position in a prior state?

Each question addresses the problem from a different perspective; as a result, a different methodology was proposed for each, as described in the corresponding sections.

The team identified rural curves of interest using the RID and requested time series data from the DAS, which provided vehicle kinematics (e.g., speed, acceleration) for those curves. A forward roadway view was also provided. Vehicle, roadway, and environmental data were extracted and used as variables in the various analyses. Eyes-off-roadway distractions and driver glance locations were reduced using the driver face and steering wheel video data at the VTTI secure data enclave.

Crash surrogates were used because crashes/near crashes had not been coded at the time this research was conducted. A number of potential crash surrogates were considered against the data available and the expected accuracy of relevant variables in the NDS data (e.g., lane position, forward radar, vehicle position). Lane offset was the best crash surrogate, but lane offset was not reliable in a number of traces. As a result, it was used for Research Questions 2 and 4, resulting in a smaller sample of data for those research questions.

Because offset was not reliable in a number of traces, it was determined that encroachments would be the best crash surrogate for Research Question 3. A right-side encroachment is defined as the right side of the vehicle crossing the right lane line, and a left-side encroachment is defined as the left side of the vehicle crossing the centerline.

Discussion and Recommendations for Countermeasures

Data from the SHRP 2 NDS and RID were used to develop relationships between driver, roadway, and environmental characteristics and risk of a roadway departure on rural two-lane curves on paved roadways.

The four research questions addressed the problem from different perspectives, and a different methodology was developed specific to each. The analytical method, data sampling and segmentation approach, general variables considered, results, and implications are discussed in the corresponding sections. In general, the research questions covered three areas: curve area of influence, lane position, and speed.

Curve Area of Influence

Identifying the curve area of interest was addressed in Research Question 1. Regression and Bayesian analyses were used to model the point (upstream of the PC) at which the driver reacts to the curve. Reaction point was inferred as the point at which speed or gas pedal position changed significantly from upstream driving.

Results showed that, depending on radius of curve, drivers begin reacting to the curve 164 m to 180 m (538.1 ft to 590.6 ft) upstream of the point of curvature. Reaction point was compared with sign placement guidelines in the *Manual on Uniform Traffic Control Devices* (Federal Highway Administration 2009). It was determined these guidelines are appropriately set based on where drivers actually react to the curve.

Research Question 1 also found that drivers begin reacting to the curve sooner for curves with larger radii than for curves with smaller radii. Drivers may not be able to gauge the sharpness of the curve, or sight distance issues may be a concern for sharper curves. This suggests that use of countermeasures, such as chevrons or RPMs, that better delineate the curve may provide better advance information for drivers. It should be noted that the model only identified where drivers reacted to the curve. This research question did not attempt to answer whether the reaction point was sufficient for drivers to successfully negotiate the curve.

Lane Position

Lane position was modeled in Research Questions 2, 3, and 4. Offset from lane position was modeled for normal driving in Research Question 2, and probability of a right-side or left-side encroachment was modeled in Research Question 3. Research Question 4 used time series data to model driver behavior at 0.1-s intervals based on driver, traffic, and roadway characteristics. However, the objective of Research Question 4 was to demonstrate the utility of the approach, and only limited data were used in the analyses.

Several driver factors are related to lane position. Research Question 2 found that offset from the center of the lane within the curve is influenced by offset upstream of the curve. Results from Research Questions 2 and 3 indicate that offset and likelihood of an encroachment are correlated to glances away from the forward roadway and glances associated with a distraction. Males are more likely to have a left-side encroachment, and younger drivers are more likely to deviate within their lane.

Research Question 2 also indicated that offset from the center of the lane varies with position with the curve. Research Questions 2 and 3 found that behavior differs when driving on the inside versus the outside of the curve (from the perspective of the driver). Because there are natural variations in position along the curve, drivers may be more vulnerable to lane departures at certain points in the curve. These results suggest that countermeasures such as rumble strips, paved shoulders, and high-friction treatments may ameliorate the consequences of variations in lane position through the curve.

Several roadway characteristics were correlated to lane position. Research Question 2 found that nighttime driving was a factor in driver lane position, with offset tending more toward the left of the lane center. Results from Research Question 3 showed a correlation between radius and the probability of encroachment, but the effect was small. These results also indicated that drivers were more likely to have a right-side encroachment on curves when an advisory sign was present but less likely when a guardrail was present along the curve. It should be noted that advisory signs and guardrails are used on certain types of curves, and, as such, either may be a surrogate for a certain type of curve.

These results suggest that presence of advisory signs do not in and of themselves mitigate roadway departures. Additionally, drivers may adjust their speed when the roadway suggests a more dangerous situation (i.e., presence of a guardrail suggests an unforgiving roadside environment). Consequently, better curve delineation may allow drivers to better gauge upcoming changes in roadway geometry, resulting in better speed selection and decreased risk of a roadway departure, and may help decrease speed. Delineation countermeasures include chevrons, the addition of reflective panels to existing chevron posts, reflective barrier delineation, RPMs, post-mounted delineators, edge lines, and wider edge lines.

Speed

Driver speed near the curve entry was modeled in Research Question 2, and probability of a driver exceeding the advisory speed (if present) or posted speed (if not present) by 8 km/h and 16 km/h (5 mph and 10 mph) was modeled in Research Question 3.

Several driver characteristics were related to speed. Both Research Questions 2 and 3 indicated that drivers traveling at higher speeds in the tangent section were also likely to speed within the curve, and younger drivers and pick-up truck/SUV drivers were more likely to speed. Both research questions also found that drivers looking away from the roadway had slower speeds within the curve.

Several roadway characteristics were correlated to speed. Results from Research Question 3 showed that drivers are slightly more likely to exceed the curve advisory/posted speed limit as radius of curve decreases, but this may be because sharper curves have lower advisory speeds. Results also showed that drivers are less likely to speed when paved shoulders or RPMs are present. The probability that a driver would exceed the posted/advisory speed by 8 km/h (5 mph) or more was higher for curves with obscured/missing edge lines.

When roadway characteristics are considered together, the results may suggest that appropriate delineation as provided by RPMs and presence of edge lines may provide cues to drivers, allowing them to better gauge the sharpness of the curve and select appropriate speeds. Delineation countermeasures include chevrons, the addition of reflective panels to existing chevron posts, reflective barrier delineation, RPMs, post-mounted delineators, edge lines, and wider edge lines.

The speed models also suggest that driver age and upstream speed have a significant impact on drivers' speed within a curve. As a result, speed management countermeasures that affect tangent speed will also decrease curve speeds. The model also indicates that speed management is appropriate to get drivers' attention before entering a curve. Countermeasures specifically targeted to reduce speed on curves include dynamic speed feedback signs, on-pavement curve warning signs, and flashing beacons.

References

AASHTO. *Driving Down Lane-Departure Crashes: A National Priority.* American Association of State Highway and Transportation Officials, Washington, D.C., 2008.

Angell, L., J. Auflick, P.A. Austria, D. Kochhar, L. Tijerina, W. Biever, T. Diptiman, J. Hogsett, and S. Kiger. *Driver Workload Metrics.* National Highway Traffic Safety Administration, U.S. Department of Transportation, 2006.

Burgett, A., and K. Gunderson. *NHTSA Research Note: Crash Prevention Boundary for Road Departure Crashes—Derivation.* National Highway Traffic Safety Administration, 2001.

Campbell, K.L. The SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety. *TR News,* No. 282, September–October 2012, pp. 30–35.

Charlton, S.G. The Role of Attention in Horizontal Curves: A Comparison of Advance Warning, Delineation, and Road Marking Treatments. *Accident Analysis and Prevention,* Vol. 39, 2007, pp. 873–885.

Charlton, S.G., and J.J. DePont. *Curve Speed Management.* Land Transport New Zealand Research Report 323. Land Transport New Zealand, Wellington, 2007.

Chin, H., S. Quek, and R. Chou. Quantitative Examination of Traffic Conflicts. In *Transportation Research Record 1376,* TRB, National Research Council, Washington, D.C., 1992, pp. 87–91.

Council, F.M. Safety Benefits of Spiral Transitions on Horizontal Curves on Two-Lane Rural Roads. In *Transportation Research Record 1635,* TRB, National Research Council, Washington, D.C., 1998, pp. 10–17.

Dingus, T., J. Hankey, J. Antin, S. Lee, E. Eichelberger, S. Klauer, J. Sudweeks, C. Gaylord, D. McGraw, S. McLaughlin, M. Mollenhauer, M. Perez, A. Petersen, A. Alden, P. Madison, T. Gordon, D. LeBlanc, H. Sardar, D. Blower, J. Sayer, M. Barnes, A. Thurtell, J. Patel, J. Campbell, C. Richard, and J. Devonshire. *Design of the In-Vehicle Driving Behavior and Crash Risk Study. Task 9: Data System Interim Report (Task 6: Driver Face and Other Video Recording and Processing).* Virginia Tech Transportation Institute, Blacksburg, 2008.

Donmez, B., L. Boyle, and J. Lee. The Impact of Driver Distraction Mitigation Strategies on Driving Performance. *Human Factors,* Vol. 48, No. 4, 2006, pp. 785–804.

Donnell, E.T., M.D. Gemar, and I. Cruzado. *Operational Effects of Wide Edge Lines Applied to Horizontal Curves on Two-Lane Rural Highways.* Pennsylvania Transportation Institute, University Park, 2006.

Farmer, C.M., and A.K. Lund. Rollover Risk of Cars and Light Trucks After Accounting for Driver and Environmental Factors. *Accident Analysis and Prevention,* Vol. 34, 2002, pp. 163–173.

Federal Highway Administration, U.S. Department of Transportation. *Roadway Departure Safety.* http://safety.fhwa.dot.gov/roadway_dept. Accessed July 2014.

Federal Highway Administration. *Manual on Uniform Traffic Control Devices.* U.S. Department of Transportation, 2009. http://mutcd.fhwa.dot.gov/pdfs/2009r1r2/pdf_index.htm. Accessed July 2014.

Fink, K.L., and R.A. Krammes. Tangent Length and Sight Distance Effects on Accident Rates at Horizontal Curves on Rural Two-Lane Highways. In *Transportation Research Record 1500,* TRB, National Research Council, Washington, D.C., 1995, pp. 162–168.

Fitzpatrick, K., A.H. Parham, and M.A. Brewer. *Treatments for Crashes on Rural Two-Lane Highways in Texas.* Report FHWA/TX-02/4048-2. Texas Transportation Institute, College Station, 2002.

Garder, P. Segment Characteristics and Severity of Head-On Crashes on Two-Lane Rural Highways in Maine. *Accident Analysis and Prevention,* Vol. 38, 2006, pp. 652–661.

Gettman, D., and L. Head. *Surrogate Safety Measures from Traffic Simulation Models.* Report FHWA-RD-03-050. Federal Highway Administration, U.S. Department of Transportation, 2003.

Glennon, J.C., T.R. Neuman, and J.E. Leisch. *Safety and Operational Considerations for Design of Rural Highway Curves.* Report FHWA/RD-86-035. Federal Highway Administration, U.S. Department of Transportation, 1985.

Hallmark, S., L. Boyle, A. Carriquiry, Y. Hsu, Y. Tian, and A. Mudgal. *SHRP 2 Report S2-S01E-RW-1: Evaluation of Data Needs, Crash Surrogates, and Analysis Methods to Address Lane Departure Research Questions Using Naturalistic Driving Study Data.* Transportation Research Board of the National Academies, Washington, D.C., 2011. http://www.trb.org/Publications/Blurbs/166050.aspx.

Luediger L., E.M. Choueiri, J.C. Hayward, and A. Paluri. Possible Design Procedure to Promote Design Consistency in Highway Geometric Design on Two-Lane Rural Roads. In *Transportation Research Record 1195,* TRB, National Research Council, Washington, D.C., 1988, pp. 111–122.

McLaughlin, S.B., J.M. Hankey, S.G. Klauer, and T.A. Dingus. *Contributing Factors to Run-Off-Road Crashes and Near-Crashes.* Report DOT HS 811 079. National Highway Traffic Safety Administration, 2009.

Miaou, S. *Estimating Roadside Encroachment Rates with the Combined Strengths of Accident and Encroachment-Based Approaches.* Report FHWA-RD-01-124. Federal Highway Administration, U.S. Department of Transportation, 2001.

Miaou, S.-P., and H. Lum. Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements. In

Transportation Research Record 1407, TRB, National Research Council, Washington, D.C., 1993, pp. 11–24.

Milton, J., and F. Mannering. The Relationship Among Highway Geometric, Traffic-Related Elements, and Motor-Vehicle Accident Frequencies. *Transportation,* Vol. 25, 1998, pp. 395–413.

Muggeo, V. M. R. R Package to Fit Regression Models with Broken-Line Relationships. *R News,* Vol. 8, No. 1, 2008, pp. 20–25.

Porter, R. J., E. T. Donnell, and K. Mahoney. Evaluation of Effects of Centerline Rumble Strips on Lateral Vehicle Placement and Speed. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1862,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 10–16.

Preston, H. Low-Cost Treatments for Horizontal Curve Safety. Presented at Federal Highway Administration webinar, 2009.

Schurr, K. S., P. T. McCoy, G. Pesti, and R. Huff. Relationship of Design, Operating, and Posted Speeds on Horizontal Curves of Rural Two-Lane Highways in Nebraska. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1796,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 18–25.

Shinar, D., E. D. McDowell, and T. H. Rockwell. Eye Movements in Curve Negotiation. *Human Factors,* Vol. 19, 1977, pp. 63–71.

Smith, S. *Digital Signal Processing: A Practical Guide for Engineers and Scientists.* Elsevier Science, Burlington, Mass., 2003.

Stodart, B. P., and E. T. Donnell. Speed and Lateral Position Models from Controlled Nighttime Driving Experiment. *ASCE Journal of Transportation Engineering,* Vol. 134, No. 11, 2008, pp. 439–449.

Taylor, M. C., A. Baruya, and J. V. Kennedy. *The Relationship Between Speed and Accidents on Rural Single-Carriageway Roads.* TRL Report TRL511. TRL, Berks, United Kingdom, 2002.

Zegeer, C. V., J. R. Stewart, F. M. Council, and D. W. Reinfurt. *Safety Effects of Geometric Improvements on Horizontal Curves,* University of North Carolina, Chapel Hill, 1991.

APPENDIX A

Methodology for Reducing Roadway Data

The methodology used to reduce various roadway data features is described in the sections below.

Kinematic Vehicle Factors

Data element: Vehicle position within its lane

Need: Lane position may be the best indicator of when a lane departure has occurred. Lane position can also be used to determine the magnitude of the lane departure in terms of departure angle from the roadway and amount that the vehicle encroaches onto the shoulder. Both can be used to set thresholds between different levels of crash surrogates.
Potential source for data element: Data can only be obtained from lane position tracking algorithms and associated data streams such as forward video.
Accuracy: Not yet available from Virginia Tech Transportation Institute (VTTI).
Resolution: 10 Hz.
Comments: The NDS DAS reports information that can be used to establish lane position. Lane-tracking units were reported as centimeters in the data dictionary, but a review of the first data set indicated this was erroneous. In a follow-up conversation with VTTI, it was determined that the units initially reported are millimeters. The following variables are used to calculate lane position (see also Figure A.1):

- Lane Position Offset (vtti.lane_distance_off_center): Distance to the left or right of the center of the lane based on machine vision.
- Lane Width (vtti.lane_width): Distance between the inside edge of the innermost lane marking to the left and right of the vehicle. Note that lane width is calculated for each 0.1-second interval and varies somewhat.
- Lane Marking, Distance, Left (vtti.left_line_right_distance): Distance from vehicle centerline to inside of left-side lane marker based on vehicle-based machine vision.
- Distance from vehicle centerline to inside of left-side lane marker based on vehicle-based machine vision.
- Lane Marking, Distance, Right (vtti.right_line_left_distance): Distance from vehicle centerline to inside of right-side lane marker based on vehicle-based machine vision.
- Lane Marking, Probability, Right (vtti.right_marker_probability): Probability that vehicle-based machine-vision lane-marking evaluation is providing correct data for the right-side lane markings. Higher values indicate greater probability.
- Lane Marking, Probability, Left (vtti.left_marker_probability): Probability that vehicle-based machine-vision lane-marking evaluation is providing correct data for the left-side lane markings.

Offset from lane center and distance from the right lane (R_D) or left lane (L_D) line are the metrics currently being used as crash surrogates. R_D and L_D are calculated as shown in Equations A.1 and A.2 (in meters).

$$L_D = -L_{CL} - T_w/2 \tag{A.1}$$

$$R_D = R_{CL} - T_w/2 \tag{A.2}$$

where

L_D = distance from left edge of vehicle to left edge of lane line; if negative, means left edge of car is to the left of the left edge line.
R_D = distance from right edge of vehicle to right edge of lane line; if negative, means right edge of car is to the right of the right edge line.
T_w = vehicle track width.

Data element: Presence and distance between subject vehicle and other vehicles

Need: Establish outcome from lane departure; used as a measure of level of service. Presence of other vehicles (opposing,

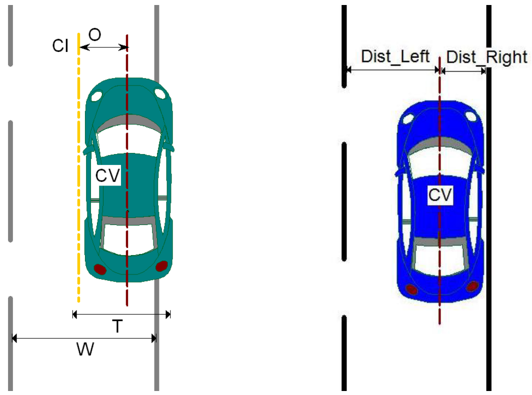


Figure A.1. Description of variables to calculate lane position.

Subject vehicle is closely following forward vehicle.

Subject vehicle is not considered to be following forward vehicle.

Source: University of Michigan Transportation Research Institute (UMTRI) Road Departure Crash Warning (RDCW) data set.

Figure A.2. Subjective measure of vehicle following.

vehicles passed) can be used to determine roadway density as an exposure method.
Source: Forward video.
Accuracy: ±3 ft (0.914 m).
Resolution: Collected as vehicle was approaching the curve.
Comments: A subjective measure of distance will be obtained from the forward video, as shown in Figure A.2, but distance cannot be determined. When a conflict occurs, distance to a forward or side vehicle will be determined from the forward or side radar. However, only vehicles within the radar range can be detected.
Coding:
Following
0: No forward vehicle present
1: Forward vehicle present but not following
2: Following closely (less than 3 seconds apart)

Roadway Factors

Data element: Lane width

Need: Independent variable in the statistical analysis; also needed to establish vehicle position within its lane.
Source: Mobile mapping when available; lane-tracking system (varies significantly over 0.1-second intervals—could use average).
Accuracy: Need to determine from mobile mapping and lane tracking.
Resolution: At curve approach, PC, apex, PT.
Comments: Lane width is measured by the DAS lane-tracking system and will be used when position within the lane is needed.
Coding: LaneWidth, reported in meters.

Data element: Shoulder width

Need: Independent variable in statistical analyses. Shoulder and median width also affect potential outcomes for lane departures.
Source: Mobile mapping data; may be available from roadway databases.
Accuracy: ±0.5 ft (0.152 m).
Resolution: At curve approach, PC, apex, PT (should be checked at several points but can be reported once).
Comments: Could not be accurately measured from aerial images and is therefore not included in initial analysis, as mobile mapping data are not available.
Coding:
Paved shoulder width
1: Less than 1 ft
2: 1 ft to less than 2 ft
3: 2 ft to less than 4 ft
4: Greater than or equal to 4 ft

Data element: Curve length and radius

Need: Independent variable in statistical analyses; may also be used to assess roll hazard.
Source: Mobile mapping, aerial imagery.
Accuracy: ±25 ft (7.62 m) for curve length and ±10% for radius.
Resolution: Once per curve.
Comments: Extracted for each direction and then averaged to find one value for each curve.

Coding:
Length of curve from PC to PT, reported in meters (Length).
Radius of curve, in meters (Radius).

Data element: Curve superelevation

Need: Independent variable in statistical analyses; may also be used to assess roll hazard.
Potential source for data element: Mobile mapping is likely the only feasible source.
Accuracy: Maximum superelevation for areas with no ice and snow is 12%; for areas with snow and ice the maximum is 8%. Given these ranges, ideal accuracy is 0.5%, but it is unknown if this accuracy can be practically measured in the field. Under normal circumstances cross slope is 1.5% to 2%. Ideally, it would be necessary to measure this variable at 0.1% accuracy to determine differences, but this may not be practical.
Resolution: Once per curve as reported by the mobile mapping.
Comments: SHRP 2 Project S04A data had both negative and positive values.
Coding: Extracted once per curve for each lane.
Superelevation, in percent (Super).

Data element: Driving direction

Need: Independent variable in statistical analyses; also important for determining the potential outcome of a noncrash lane departure.
Source: Aerial imagery and forward view.
Accuracy: NA.
Resolution: Should be indicated once per curve.
Comments: None.
Coding:
Direction of travel (Cardinal)
0: N/S
1: E/W
2: NE/SW
3: NW/SE

Direction of curve from perspective of driver (Direction)
0: Outside/left-hand
1: Inside/right-hand

Data element: Distance to upstream curve, distance to downstream curve from perspective of driver (meters)

Need: Drivers may negotiate curves differently if they have traveled for some distance between curves instead of having negotiated a series of curves. Also used as an independent variable in statistical analyses.
Source: Aerial imagery.
Accuracy: ±25 ft (7.62 m).
Resolution: Upstream and downstream per curve.
Comments: None.
Coding:
Distance to upstream curve from perspective of driver, in meters (DistUP).
Distance to downstream curve from perspective of driver, in meters (DistDown).

Curve type
0: Individual curve
1: S-curve (less than 600 ft between subsequent curves)
2: Compound curve (0 ft between 2; the PT and PC of subsequent curves in the same direction)

Data element: Speed limit, curve advisory, chevrons, and W1-6 signs

Need: Independent variable in statistical analyses.
Source:

- Speed limit and curve advisory speed limit from mobile mapping.
- Forward video/Google/forward view mobile mapping for remaining.

Accuracy: The general location of the sign or an indication that the sign is present is adequate. For instance, it would be important to know the number and type of chevrons that were present on a curve, but it is not necessary to know exactly where each sign is located. It is also assumed that all signs are compliant with National Cooperative Highway Research Program (NCHRP) 350 so that they would not need to be considered as strikable fixed objects when determining the outcome of a lane-departure event. A sign located using a standard GPS with accuracy of ±6.6 ft (2 m) would be adequate.
Resolution: As they occur.
Comments: None.
Coding:
Tangent speed limit (SpdLimit), in miles per hour.
Advisory speed (Advisory), in miles per hour, or 999 if no advisory speed limit exists.

Presence of chevrons (Chevrons)
0: Not present
1: Present

Presence of curve advisory sign
0: Not present
1: Present

Presence of W1-6 sign
0: Not present
1: Present

Data element: Number of driveways or other access points

Need: Traffic entering and exiting the traffic stream can affect vehicle operation. This traffic would be included as an independent variable in statistical analyses.
Source: Aerial imagery and forward imagery.
Accuracy: NA.
Resolution: Number in the upstream, curve, and downstream.
Comments: Four-way intersections counted as one cross street.
Coding: Number of driveways at approach, within curve, at exit.
Cross streets (CrossStreets), in points per section through length of curve and tangents.
Driveways (Dwys), in driveways per section through length of curve and tangents.

Data element: Presence of edge or centerline rumble strips

Need: Independent variable in statistical analyses; also needed to establish outcome of lane departure.
Source: Forward video and Google Street View.
Accuracy: NA.
Resolution: Curve approach and in curve.
Comments on extracting data from existing data sets: Only presence of rumble strip could be extracted, not distance from road.
Coding:
Type of rumble strip (RS)
0: No rumble strip present
1: Edge line rumble strips only (see Figure A.3)
2: Centerline rumble strips only
3: Centerline and edge line rumble strips

Data element: Roadway delineation (presence of lane lines or other on-roadway markings)

Need: Critical for lane position tracking software; would be included as an independent variable in statistical analyses.

Source: DAS forward imagery.

Figure A.3. Presence of edge-line-only rumble strips.

Source: Forward view.
Desired accuracy: Data is a quantitative estimate of visibility of markings.
Resolution: Once per mile or as situation changes.
Comments: This element needs to be current to driving situation and can only be extracted from forward imagery. This information could be obtained from the UMTRI data set but was more difficult with the VTTI data set due to image resolution.
Coding:
Presence of raised pavement markings (RPMs)
0: Not present
1: Present

Roadway delineation (Delineation)
0: Highly visible
1: Visible
2: Obscured
3: Not present

Figure A.4 shows an example of a subjective measure.

Data element: Roadway furniture

Need: Necessary to determine how roadside makeup affects driving; also how roadway furniture may affect the severity of a lane-departure crash.
Source: Forward view.
Accuracy: NA.
Resolution: Once per curve just upstream of PC looking at curve ahead for roadway furniture rating; once per curve at any location for presence of guardrail.
Coding:
Presence of guardrail
0: Not present
1: Present

Roadway furniture
1: Little to no roadway furniture
2: Moderate roadway furniture
3: Large amount of roadway furniture

Figure A.5 shows an example of a subjective measure.

Data element: Sight distance

Need: The distance at which the curve is first visible will have an effect on where a driver reacts to the curve and could play a role in lane departures.
Source: Forward view and time series data.
Accuracy: NA.
Resolution: Once per direction per curve.

Pavement markings indicated as "highly visible."

Pavement markings indicated as "visible."

Right pavement markings indicated as "obscured."

Source: Forward video and UMTRI RDCW data set.

Figure A.4. Subjective measure of lane marking condition using forward imagery.

Comments: This was calculated once per curve using the best forward video available. At times, night was the only condition to assess sight distance of the curve. Timestamp at which curve could first be seen was recorded and then used to find corresponding distance upstream in time series data.
Coding: Distance in meters to PC.

Environmental Factors

This section summarizes environmental factors necessary to address lane-departure research questions, indicates potential sources in the existing data sets, suggests accuracy and frequency needs, and includes comments about the accuracy and availability in the existing data sets.

Data element: Roadway surface condition (presence of roadway irregularities such as potholes)

Need: Independent variable in statistical analyses; may also affect potential outcome of lane departure.
Source: Forward or other outward facing video, status and frequency of wiper blades, outside temperature if available, roadway weather information system (RWIS) data if archived.
Accuracy: Measure is subjective and therefore not applicable.
Resolution: At curve approach, in curve.
Comments: None.
Coding:
Roadway surface condition (PaveCnd)
0: Normal surface condition, no obvious damage present
1: Moderate damage
2: Severe damage, presence of potholes

Figure A.6 shows an example of a subjective measure.

Data element: Environmental conditions such as raining, snowing, cloudy, clear (may not correspond to roadway surface condition)

Need: Independent variable in statistical analyses; may affect sight distance and is related to visibility.
Source: Forward imagery or archived weather information, ambient temperature probe.
Accuracy: Subjective measure.
Resolution: Once per vehicle trace.
Comments: A general assessment of environmental conditions can be obtained from the forward video (Figure A.7). Even with wiper position, it is difficult to tell how heavy

Little to no roadway furniture.

Moderate roadway furniture.

Large amount of roadway furniture.

Source: DAS forward imagery.

Figure A.5. Subjective measure of roadway furniture.

rainfall is. Archived weather information can provide general information for an area but cannot tell the exact environmental conditions for the location where the subject vehicle is located.

Coding:

Roadway surface condition (Surface)
0: Dry pavement surface
1: Pavement wet but not currently raining
2: Wet and light rain
3: Wet and heavy rain
4: Snow present but road is bare
5: Snow along road edge and/or centerline
6: Light snow on roadway surface
7: Roadway surface covered

Data element: Ambient lighting

Need: Independent variable in statistical analyses.
Source: Derived from sun angle, twilight, and forward view.
Accuracy: Subjective measures.
Resolution: Once per trace or as conditions change.
Comments: A relative estimate of ambient lighting can be obtained in most cases from the forward imagery. The limitations are that it was difficult during high cloud cover or low visibility to subjectively estimate ambient lighting.
Coding:
Ambient lighting (Lighting) time of day and lighting
0: Daytime
1: Dawn/dusk

2: Nighttime, no lighting
3: Nighttime, lighting present

Data element: Visibility

Need: Independent variable in statistical analyses; serves as a measure of sight distance and can also indicate surface conditions.
Source: Forward view is the only reasonable data source.
Accuracy: Subjective variable.
Resolution: Once per trace.
Comments: This element is available from forward imagery. In some cases it may be difficult to tell whether visibility or image resolution causes securement, as shown in Figure A.8. The source of decreased visibility could not be determined. Low visibility is shown in Figure A.9, but it is unknown if the source is fog, smoke, or dust.
Coding:
Visibility
0: Clear
1: Reduced visibility
2: Low visibility

Pavement condition indicated as "normal."

Pavement condition indicated as "moderate."

Source: DAS forward imagery.

Figure A.6. Subjective measure of roadway pavement surface condition using forward imagery.

Pavement surface condition: snow present but roadway bare.

Pavement surface condition: wet but amount of water cannot be determined.

Surface irregularities.

Source: UMTRI RDCW data set.

Figure A.7. Pavement surface condition from forward imagery.

Source: DAS forward imagery.

Figure A.8. Reduced visibility may be due to sun angle or image resolution.

Source: DAS forward imagery.

Figure A.9. Low visibility appears to be due to fog.

Exposure Factors

This section summarizes exposure factors necessary to address lane-departure research questions, indicates potential sources in the existing data sets, suggests accuracy and frequency needs, and includes comments about the accuracy and availability in the existing data sets.

Data element: Density

Need: Exposure measure.
Source: Forward video.
Accuracy: NA.
Resolution: Number of vehicles on approach, within curve, at exit.
Comments: The number of oncoming vehicles, vehicles passed by the subject vehicle, or vehicles that the subject vehicle passes can be counted using the forward and side imagery. Density can be calculated knowing the number of vehicles encountered over a specific distance. Density is a good measure of roadway level of service. However, counting vehicles in the forward or side imagery is time-consuming.
Coding: Number of vehicles passing subject vehicle during period (Density), in vehicles per meter, calculated through curve.

APPENDIX B

Data Reduction Method for Coding Driver Glance Location and Distraction

Additional coding of the video data for each of the rural road curve segments was completed by the University of Iowa. This was done to collect the following information:

- Passenger presence and environmental conditions;
- Eyeglance location, frequency, and duration; and
- Driver distraction.

Because of the identifiable nature of the video data, all coding was completed at Virginia Tech Transportation Institute's (VTTI) secure data enclave. VTTI has many procedures in place to ensure that the SHRP 2 data are protected and only used for the purpose that was specified in the data plan. Specifically, all researchers who enter the enclave are required to use a passcode and leave all electronics outside, all materials are examined when researchers leave the enclave to ensure that data are not removed, and a proctor is scheduled to be in the room whenever a researcher is present.

Initial video coding was done by examining the occupancy snapshot available for a particular event. From that frame, the research team coded front and rear passenger presence as well as environmental conditions (e.g., light, weather, road surface). For some events, the quality of the video or the size and position of the passenger made this difficult. With only a single frame, the reviewers cannot see movement that could be attributed to passengers. The lack of audio excludes another way of identifying passenger presence. On several occasions it was necessary to view additional snapshots that might have occurred outside the event but during the same drive in order to see if different lighting conditions or passenger position within the vehicle would make that information obtainable.

Eyeglance Data

Driver eyeglance data were coded beginning at the straightaway leading up to the curve and ending at a certain point beyond the curve. These data were coded frame by frame at approximately 10 Hz using Virginia Tech's Hawk-I software. Table B.1 shows the eyeglance locations and the rules associated with how the glances were defined and applied.

Eyeglance data coding under naturalistic driving conditions was challenging for a number of reasons:

- Bright sunlight caused the camera to "wash out" the entire face, especially at certain times of the day when the sunlight was more direct. External light sources at night, such as street lights, created the same effect.
- Night videos had a grainy quality, making it difficult to discern the driver's eye from the rest of the view. It was thought that this might have been caused by the artificial light created by the infrared light on the camera.
- Many drivers wear sunglasses, as well as prescription glasses, both of which create problems associated with glare.

These problems were not unexpected and do not make the videos as a whole uncodable. In many cases when the driver's eye is not visible, a dark spot indicating the pupil can be seen and used for coding. Head movements were also used to aid in the coding of some eyeglances. And glare and other problems associated with sunlight change as the direction of the vehicle changes.

One challenge the research team encountered that was unexpected was the discontinuous camera view. As researchers coded, they found that the camera zoomed in and out to get the best possible view of the driver's eyes. Therefore, the distance of the eyes from the camera varied among drivers and even within events. While this function was intended to be helpful, in many cases, it did not improve the view and in fact made it harder to remain consistent in the coding.

Driver Distraction

Visual distractions are those that cause the driver to take their eyes off the roadway. Table B.2 lists the distractions coded for this study and some of the eyeglance locations that might have been associated with that distraction. Pairing this with

Table B.1. Eyeglance Coding Rules

Location of Eyeglance	Coding Rule
Forward	Gazes to the center, left, or right that involve little or no head movement and appear to be mostly directed to the left or right portions of the windshield should be coded as "Forward."
Center console	Eyes move slightly down and to the right. There is little or no head movement (e.g., HVAC, radio).
Steering wheel	Eyes move down slightly. There is little or no head movement (e.g., speedometer, fuel gauge, cruise control).
Down	Draw an imaginary horizontal line in the middle of the steering wheel. If a gaze is directed above the line it should be coded as "Steering wheel" or "Center console." If it is below that line, it should be coded as "Down." Some head movement is associated with a "Down" glance (e.g., looking at something in lap or on floor).
Up	Eye movement to the upper-left or upper-central portion of the windshield should be coded as "Up." This glance is rare and is usually associated with the visor or sunroof, if present.
Left	Any gazes to the left of the A-pillar should be coded as "Left" whether the driver is looking at the left mirror or out the driver's side window.
Right	Any gazes that involve *both* eye and head movement to the right should be coded as "Right" whether the driver is looking at the right mirror, glove box, front-seated passenger, or out the passenger's side window.
Rearview mirror	Eye movements up and to the right with a slight head movement should be coded as "Rearview mirror." These include scanning the roadway behind the vehicle, as well as glances to the rear-seated passengers.
Over the shoulder	Any glance over the left or right shoulder of the driver. This movement requires the driver's eyes to pass the B-pillar.
Other	Blinks, squints, or closed eyes that last more than 10 frames. Any blinks, squints, or closed eyes less than that should be disregarded.
Missing	Code as "Missing" if • The eyes are obscured or obstructed for more than 10 frames. • The video freezes or video signal is dropped. • The locus of gaze cannot be inferred because of glare, excessive head movement, or camera location.

Table B.2. Potential Distractions Associated with Eyeglances

Distraction	Probable Glance Locations	Situation
Passenger	Right (front-seated passenger), rearview mirror, or over the shoulder (rear-seated passenger)	A glance associated with a front- or rear-seated passenger with indication of a conversation or other distracting activity. The glance location depends on the seating position of the passenger.
Route planning (locating, viewing, or operating)	Steering wheel, down, center console	A glance associated with the actions performed during the use of a paper map or in-vehicle navigation system. The glance location depends on where the driver holds the instrument while looking at it.
Moving or dropped object in vehicle	Down	A glance associated with the driver reaching for something in the vehicle. The glance location depends on the location of the object.
Animal/insect in vehicle	All locations are possible	A glance associated with the driver being preoccupied by the presence of an animal/insect and taking action to remedy the distraction. The mere presence is not to be coded as a distraction. The glance location depends on where the animal/insect is located in the vehicle.
Cell phone (locating, viewing, operating)	Steering wheel, down, center console	A glance associated with the actions performed during cell phone use. The glance location depends on where the driver holds the phone while looking at it.
IPod/MP3 (locating, viewing, operating)	Steering wheel, down, center console	A glance associated with the actions performed during the use of an in-vehicle entertainment system. The glance location depends on the location of the device.
In-vehicle controls	Center console, steering wheel, down	A glance associated with the actions performed using the in-vehicle controls (e.g., HVAC, radio, CD player, wipers, windows, door locks). The glance location depends on the control being activated.
Drinking/eating	Steering wheel, down	A glance associated with locating/adjusting food item or drink container. The glance location depends on where the driver is holding the food/drink.

(continued on next page)

Table B.2. Potential Distractions Associated with Eyeglances (continued)

Distraction	Probable Glance Locations	Situation
Smoking	Steering wheel, down, center console, left	A glance associated with locating, lighting, smoking, or disposing of ashes. The glance location depends on where the driver holds the cigarette and where he or she discards the ashes.
Personal hygiene	Up, rearview mirror, steering wheel, down	A glance associated with the driver performing an action related to personal hygiene (e.g., fixing hair, applying makeup, blowing nose). The glance location depends on the activity the driver is performing.
Other task	Any are possible	A glance not fitting another category (make a note if used).

the frame-by-frame eyeglance data allowed researchers to determine not only the duration of the glance but the cause.

Manual distractions are those that require the driver to take a hand off the steering wheel to perform a task unrelated to driving the vehicle. In most cases, these distractions were identifiable and codable using the video data. These distractions include drivers rubbing their nose, twirling their hair, or holding a phone to their ear. In most cases, drivers performed these actions without ever taking their eyes off the road. Time constraints did not allow researchers to code these data frame by frame. However, they were noted in the data file.

Cognitive distractions, in which drivers' attention shifts away from the task of driving, are not easily coded using naturalistic driving data. For this particular study, some of the difficulties stemmed from not having sound or a view of the passengers. Without these, it was not possible to determine whether conversations (phone or personal) were occurring in the vehicle. Even if the researcher could see that the driver's mouth was moving, it was not possible to discern whether he or she was singing, talking to himself or herself, conversing with a passenger, or using a hands-free phone.

APPENDIX C

Curve Area of Influence Model Results

Table C.1. Model Output for Speed Change Points

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
IN13Ain	1910	31.4746	0.0137	–0.009	–271.5092	<0.0001
IN13Ain	1910	25.215	7.00E-04	–0.0027	–110.6764	0.02757711
IN13Aout	1910	23.7545	–0.0026	0.0063	–286.1063	<0.0001
IN13Aout	1910	23.9949	–0.005	0.0126	–126.195	<0.0001
IN13Aout	1910	27.4545	0.0011	–0.0041	–323.8188	<0.0001
IN13Aout	1910	24.9573	–2.00E-04	0.0038	–243.2865	<0.0001
IN13Bin	3862	24.9795	–1.00E-04	–0.0027	–184.6962	<0.0001
IN13Bin	3862	25.6622	5.00E-04	–0.0059	–98.4718	<0.0001
IN13Bin	3862	26.85	0.0014	–0.0068	–76.0516	<0.0001
IN13Bin	3862	25.2465	–9.00E-04	–0.002	–185.9717	<0.0001
IN1Ain	3977	23.9963	0.0033	–0.0083	–96.2083	<0.0001
IN1Ain	3977	0.2778	0	0	–408.4287	<0.0001
IN1Bout	1196	21.066	–0.0066	0.0141	–135.2823	<0.0001
IN3Aout	7106	26.3457	–3.00E-04	–0.0855	–131.2403	0
IN3Aout	7106	24.8481	–0.0084	0.0084	–84.6043	<0.0001
IN3Aout	7106	24.7021	0.0062	–0.0218	–62.822	<0.0001
IN3Bout	5994	26.6667	0	–0.2083	–2.6667	0
IN3Bout	5994	28.2097	0.0048	–0.0099	–115.7904	<0.0001
IN3Bout	5994	26.42	0.0038	–0.0042	–75.734	0.01257546
IN3Cin	1950	20.6077	–0.0134	0.017	–266.528	<0.0001
IN3Cin	1950	–88.3059	–0.2809	0.2799	–403.1271	0.23788383
IN3Cin	1950	25.2396	0.0059	–0.0062	–206.1128	<0.0001
IN44Ain	2703	28.3537	0.0058	–0.0127	–276.1865	<0.0001
IN44Ain	2703	25.8797	0.0041	–0.012	–58.0099	<0.0001
IN44Ain	2703	26.8077	–3.00E-04	–0.0126	–76.1235	<0.0001
IN44Cout	2181	23.3964	–0.0096	0.0157	–192.072	0
IN44Cout	2181	24.3886	–0.0052	0.0061	–218.0717	<0.0001

(continued on next page)

Table C.1. Model Output for Speed Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
IN44Cout	2181	26.7444	7.00E-04	–0.0068	–230.1197	<0.0001
IN44Cout	2181	25.6731	–0.009	0.0067	–85.4777	0
IN44Dout	1527	24.5603	–0.0049	–0.0222	–147.004	0
IN44Dout	1527	27.132	0.0038	–0.0173	–192.4604	<0.0001
IN44Dout	1527	26.748	–1.00E-04	–0.0325	–61.5033	<0.0001
IN44Fout	1490	24.6346	0.0013	–0.0069	–234.6482	<0.0001
IN44Fout	1490	25.2072	0.0056	–0.0079	–105.4487	<0.0001
IN44Fout	1490	25.484	0.003	–0.0119	–63.7031	<0.0001
IN44Fout	1490	25.7215	–0.0029	–0.0061	–168.0645	<0.0001
IN44Gin	5142	21.1746	0.005	–0.0135	–99.9366	<0.0001
IN44Gin	5142	26.4831	0	–0.0057	–80.0375	<0.0001
IN44Gin	5142	27.8004	0.0041	–0.0101	–298.2119	<0.0001
IN44Iout	2428	23.2586	–0.008	0.0079	–265.9376	<0.0001
IN44Iout	2428	19.6469	–0.0126	0.0215	–271.0077	<0.0001
IN44Iout	2428	23.7188	–0.0046	0.0152	–279.501	<0.0001
IN44Iout	2428	27.2782	0.0039	0.0013	–100	<0.0001
IN44Jout	1971	15.5089	–0.0138	0.0495	–84.071	<0.0001
IN44Jout	1971	24.3301	–0.0057	0.0044	–170.6553	<0.0001
IN44Jout	1971	24.0206	–0.0058	0.0107	–117.4916	<0.0001
IN44Kout	2051	28.3407	–0.0027	–0.0097	–268.0207	<0.0001
IN44Kout	2051	25.8547	–0.0015	–0.0067	–159.2281	<0.0001
IN44Kout	2051	21.3862	–0.0167	0.0175	–269.2508	<0.0001
IN44Kout	2051	25.4666	–0.0038	–0.0074	–117.0499	<0.0001
IN7Aout	1994	26.008	0.0052	–0.0229	–66.2461	<0.0001
IN7Aout	1994	31.7183	0.0113	–0.0148	–239.1769	<0.0001
NC16Ain	1551	22.3767	0.0035	–0.0077	–90.8977	<0.0001
NC16Ain	1551	19.7161	–2.00E-04	–0.0066	–42.009	<0.0001
NC16Cout	2053	15.531	–0.0101	0.0209	–264.3996	0
NC16Cout	2053	20.4671	0.0031	–0.3243	–2.4041	0
NC16Din	2117	24.7885	0.0089	–0.011	–388.8707	<0.0001
NC16Din	2117	22.1162	0.0044	–0.0073	–230.1436	<0.0001
NC3Ain	1129	24.6616	–0.0049	0.02	–78.9852	<0.0001
NC3Ain	1129	19.6051	–0.0093	0.0067	–359.9583	0.49582818
NC3Ain	1129	26.873	0.0041	–0.0099	–101.9631	<0.0001
NC3Ain	1129	27.5885	0.001	–0.0456	–13.7497	0.00010601
NC3Ain	1129	31.115	0.0135	–0.0169	–306.7835	<0.0001
NC3Ain	1129	27.7387	0.0076	–0.0117	–155.9568	<0.0001
NC3Aout	1129	28.384	–0.0014	–0.0052	–121.5323	<0.0001
NC3Aout	1129	28.0548	0.0013	–0.0148	–193.4051	<0.0001

(continued on next page)

Table C.1. Model Output for Speed Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
NC3Aout	1129	28.9097	0.0043	–0.0104	–226.9498	<0.0001
NC3Aout	1129	30.4472	0.0054	–0.0301	–66.6559	<0.0001
NC3Aout	1129	27.8477	0.0022	–0.0198	–68.3807	<0.0001
NC3Aout	1129	29.4495	0.0047	–0.0179	–84.8276	<0.0001
NC7Ain	1994	25.4649	0.0019	–0.0104	–106.7491	<0.0001
NC7Ain	1994	25.3715	0.0031	–0.0141	–159.2339	<0.0001
NC7Cin	723	24.5092	0.0162	–0.0242	–103.2604	<0.0001
NC7Cin	723	21.0862	–0.0022	–0.0273	–36.3002	<0.0001
NC7Eout	2153	22.9851	0.0067	–0.0383	–23.8492	<0.0001
NC7Eout	2153	24.0844	0.0034	–0.011	–75.2521	<0.0001
NC7Fin	963	27.7699	0.0096	–0.0148	–245.5055	<0.0001
NC7Fin	963	25.666	0.0061	–0.0189	–278.1089	<0.0001
NY18Ain	1718	21.7288	–0.0014	0.0112	–209.5014	<0.0001
NY18Ain	1718	31.2833	0.0204	–0.025	–306.2746	0
NY18Cin	2892	23.5163	–0.0054	0.0108	–252.1221	0
NY18Cin	2892	24.6396	–0.0078	0.0095	–223.4709	<0.0001
NY18Cin	2892	24.1158	–0.0013	0.0145	–160.5519	0
NY18Cin	2892	21.2281	–0.0106	0.0121	–292.4173	0
NY18Cin	2892	24.7382	–0.0011	–0.0021	–44.7614	<0.0001
NY18Cin	2892	24.5923	–1.00E-04	–0.0069	–64.6186	<0.0001
NY18Cin	2892	22.8488	–0.0031	0.0068	–260.3266	0
NY18Cin	2892	21.7511	–0.0036	0.0092	–101.4367	<0.0001
NY23Ain	6291	9.3436	–0.044	0.1136	–166.1236	<0.0001
NY23Ain	6291	23.5102	0.0028	–0.0319	–38.045	<0.0001
NY23Aout	6291	28.6649	0.008	–0.015	–235.6528	<0.0001
NY23Aout	6291	23.7078	–0.0023	0.0045	–242.8928	<0.0001
NY52Aout	1876	21.3175	–0.0082	0.0065	–281.4522	<0.0001
NY52Aout	1876	24.4018	–0.011	0.0181	–414.73	<0.0001
NY52Cin	1523	25.6515	0.0063	–0.011	–46.9281	<0.0001
NY52Cin	1523	26.7531	0.0096	–0.0325	–54.2924	<0.0001
NY52Cout	1523	29.8229	0.0072	–0.0072	–247.4987	<0.0001
NY52Din	1378	28.7769	0.0034	–0.0034	–48.5786	0.0005265
NY52Dout	1378	24.0518	–3.00E-04	0.0029	–128.0449	<0.0001
NY52Dout	1378	27.5435	0.0088	–0.7233	–3.1323	<0.0001
NY60Ain	1218	22.4699	0.0067	–0.0129	–261.8548	<0.0001
NY60Ain	1218	18.5039	–0.0018	–0.0055	–222.1382	<0.0001
NY62Aout	1926	31.6968	0.0166	–0.0172	–282.006	0
NY62Aout	1926	25.229	0.0055	–0.017	–239.5048	0
NY63Aout	1357	21.361	0.0044	–0.0147	–179.1604	<0.0001

(continued on next page)

Table C.1. Model Output for Speed Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
NY63Aout	1357	22.5121	–9.00E-04	–0.0066	–233.9469	<0.0001
NY63Aout	1357	20.4114	–0.0027	0.0115	–53.5696	<0.0001
NY63Aout	1357	26.7932	0.0161	–0.0193	–317.9716	<0.0001
NY63Aout	1357	22.8013	0.0064	–0.0059	–246.838	<0.0001
NY63Aout	1357	18.757	–0.0011	0.0073	–117.4594	0.00010972
NY65Ain	828	16.6772	–0.014	0.011	–218.7262	<0.0001
NY65Ain	828	21.3056	–3.00E-04	0.009	–46.6096	<0.0001
NY65Aout	828	28.5619	0.0152	–0.0245	–306.3977	<0.0001
NY65Aout	828	21.645	5.00E-04	0.0011	–299.0026	0.43003242
NY65Aout	828	20.8963	1.00E-04	0.0042	–75.7239	<0.0001
NY65Bin	1175	21.9469	–0.0015	–0.0058	–75.7584	<0.0001
NY65Bin	1175	18.0093	–0.0108	0.0152	–216.5231	<0.0001
NY65Bin	1175	22.2093	0.0067	–0.0176	–35.1011	<0.0001
NY65Bout	1175	18.1746	–0.0017	0.0092	–91.2507	<0.0001
NY65Bout	1175	18.1807	–0.0037	0.0116	–105.0212	<0.0001
NY67Aout	117	19.169	–0.0015	–0.0674	–86.9569	<0.0001
NY67Aout	117	20.0644	–0.0023	–0.0527	–102.7345	0
NY67Aout	117	19.0129	–0.0023	–0.0609	–94.651	<0.0001
NY67Aout	117	19.9117	–0.0049	–0.0523	–100.2466	0
NY67Aout	117	20.4601	–0.004	–0.1246	–57.6928	0
NY67Aout	117	19.0357	7.00E-04	–0.0612	–149.5927	<0.0001
NY67Aout	117	18.0293	–0.0041	–0.0723	–85.5433	<0.0001

Table C.2. Model Output for Pedal Change Points

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
IN13Ain	1910	16.2573	–0.0047	–0.045	–83.7498	0.00060469
IN13Ain	1910	19.1604	–0.0013	0.0246	–102.4196	0.03682925
IN13Aout	1910	34.43	0.0333	–0.1058	–202.66	<0.0001
IN13Aout	1910	27.3095	0.042	–0.0705	–271.0789	<0.0001
IN13Aout	1910	–0.6262	–0.0614	0.0982	–254.6267	0
IN13Aout	1910	34.6946	0.0341	–0.1178	–180.0157	0
IN13Bin	3862	19.3653	0.0019	0.0607	–133.6037	0
IN13Bin	3862	8.8224	–4.00E-04	0.0227	–302.1735	<0.0001
IN13Bin	3862	12.5853	–0.08	0.1146	–81.0016	<0.0001
IN13Bin	3862	21.6542	0.0081	0.0626	–115.9072	<0.0001
IN1Ain	3977	7.1324	–0.0376	0.0936	–36.267	<0.0001
IN1Ain	3977	46.1968	0.0575	–0.1781	–202.7759	<0.0001
IN1Bout	1196	27.2941	0.0417	–0.1661	–88.1225	<0.0001

(continued on next page)

Table C.2. Model Output for Pedal Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
IN3Aout	7106	0.7931	–0.0649	0.8892	–13.5572	<0.0001
IN3Aout	7106	–40.011	–0.1912	0.206	–281.5269	<0.0001
IN3Aout	7106	79.6819	0.1345	–0.1912	–308.8656	<0.0001
IN3Bout	5994	14.7636	0.0137	–0.0633	–106.6624	<0.0001
IN3Bout	5994	8.3226	–0.0381	0.1261	–70.6715	<0.0001
IN3Bout	5994	26.2856	0.0094	–0.0114	–151.2468	<0.0001
IN3Cin	1950	14.7162	0.0178	–0.0445	–166.0595	<0.0001
IN3Cin	1950	537.62	1.2923	–1.3111	–404.0769	<0.0001
IN3Cin	1950	37.7549	0.0363	–0.069	–296.7558	<0.0001
IN44Ain	2703	12.9683	–0.0645	0.1067	–181.1062	<0.0001
IN44Ain	2703	60.5642	0.076	–0.1224	–257.0432	<0.0001
IN44Ain	2703	30.6345	–0.011	–0.1332	–68.529	<0.0001
IN44Cout	2181	24.0248	–0.02	0.0434	–275.9222	0.0003697
IN44Cout	2181	23.8753	–0.0018	0.6534	–18.0016	<0.0001
IN44Cout	2181	130.6829	0.297	–0.342	–349.2169	<0.0001
IN44Cout	2181	14.2234	–0.0412	0.052	–179.9629	<0.0001
IN44Dout	1527	53.4996	0.0721	–0.1598	–219.2371	<0.0001
IN44Dout	1527	111.2334	0.2086	–0.3412	–259.2163	<0.0001
IN44Dout	1527	49.3589	0.0343	–0.2503	–140.8022	<0.0001
IN44Fout	1490	20.8767	–0.0398	0.0754	–160.7445	<0.0001
IN44Fout	1490	49.8587	0.0753	–0.1029	–265.3079	<0.0001
IN44Fout	1490	21.5665	–0.003	–0.0454	–131.8775	<0.0001
IN44Fout	1490	19.5188	–0.0216	0.027	–131.8601	<0.0001
IN44Gin	5142	35.376	0.0223	–0.0882	–222.0049	<0.0001
IN44Gin	5142	29.7204	–0.0039	–0.0401	–182.4526	0.0021038
IN44Gin	5142	12.9394	–0.0506	0.0368	–179.0259	0.00012747
IN44Iout	2428	–125.095	–0.3834	0.4156	–377.9875	<0.0001
IN44Iout	2428	25.0592	0.0199	0.3407	–22.5356	<0.0001
IN44Iout	2428	35.9075	0.0676	–0.0925	–177.8053	<0.0001
IN44Iout	2428	32.3482	0.0341	–0.0324	–185.3678	<0.0001
IN44Jout	1971	–4.1763	–0.1128	0.243	–183.2947	<0.0001
IN44Jout	1971	32.0948	–0.0097	–0.0553	–172.3349	<0.0001
IN44Jout	1971	41.634	0.0231	–0.0729	–254.8227	<0.0001
IN44Kout	2051	24.3978	–0.0155	0.0431	–224.8909	<0.0001
IN44Kout	2051	52.5632	0.0889	–0.0859	–303.1924	<0.0001
IN44Kout	2051	43.4282	0.0958	–0.0759	–175.2404	<0.0001
IN44Kout	2051	30.4323	0.0186	–0.0532	–132.7226	<0.0001
IN7Aout	1994	11.1807	0.0132	–0.0729	–95.7053	<0.0001
IN7Aout	1994	2.5765	–0.0628	0.0815	–180.6707	<0.0001

(continued on next page)

Table C.2. Model Output for Pedal Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
NC16Ain	1551	23.5	0	0.1159	–69.2133	<0.0001
NC16Ain	1551	43.5073	0.2449	–0.215	–75.4886	<0.0001
NC16Cout	2053	–309.0423	–0.7841	0.8808	–391.1233	<0.0001
NC16Cout	2053	11.7584	–0.013	0.1921	–48.3185	<0.0001
NC16Din	2117	11.1766	–0.0273	0.0447	–409.3794	<0.0001
NC16Din	2117	30.0628	0.0095	–0.2398	–71.3046	<0.0001
NC3Ain	1129	–32.209	–0.1805	0.3179	–232.3429	<0.0001
NC3Ain	1129	11.9322	0.0188	–0.1412	–61.2861	<0.0001
NC3Ain	1129	21.9174	0.0251	–0.0964	–163.3337	<0.0001
NC3Ain	1129	15.4312	0.0128	–0.3152	–35.7398	0.02054192
NC3Ain	1129	–32.6403	–0.1551	0.1419	–290.0104	<0.0001
NC3Ain	1129	14.0962	–0.009	–0.1668	–67.2934	0.00199844
NC3Aout	1129	20.0111	0.0065	0.1578	–109.63	<0.0001
NC3Aout	1129	1.4423	–0.0315	0.1909	–120.2512	<0.0001
NC3Aout	1129	8.4504	–0.0085	0.3792	–56.5933	<0.0001
NC3Aout	1129	29.9055	0.0432	–0.1855	–133.5346	<0.0001
NC3Aout	1129	–30.7789	–0.1224	0.1505	–321.0941	<0.0001
NC3Aout	1129	9.7035	–0.0142	0.3241	–52.322	<0.0001
NC7Ain	1994	26.0824	0.0146	–0.0962	–86.5459	<0.0001
NC7Ain	1994	21.2397	–0.0057	–0.0085	–141.8584	0.07829868
NC7Cin	723	15.9173	–0.0737	0.1867	–32.0757	0.00014077
NC7Cin	723	30.6289	0.0455	–0.1286	–126.5949	<0.0001
NC7Eout	2153	56.6284	0.0976	–0.1007	–332.1445	0.00036162
NC7Eout	2153	19.5064	–0.0098	0.139	–25.5921	0.00927958
NC7Fin	963	35.9203	0.0469	–0.107	–169.5773	<0.0001
NC7Fin	963	24.3556	0.0149	–0.0988	–80.2646	<0.0001
NY18Ain	1718	61.9878	0.1793	–0.2703	–143.3261	<0.0001
NY18Ain	1718	20.862	–0.0043	0.0266	–292.4055	<0.0001
NY18Cin	2892	37.0391	0.068	–0.0924	–205.5335	<0.0001
NY18Cin	2892	–48.2932	–0.1681	0.1868	–366.4744	<0.0001
NY18Cin	2892	41.4428	0.0857	–0.0857	–51.8138	<0.0001
NY18Cin	2892	70.4704	0.183	–0.1682	–291.5517	<0.0001
NY18Cin	2892	28.5348	0.0368	–0.0562	–313.1313	<0.0001
NY18Cin	2892	77.3089	0.1682	–0.2293	–295.2785	<0.0001
NY18Cin	2892	32.7137	0.0489	–0.0536	–225.2202	<0.0001
NY18Cin	2892	–1.4477	–0.0479	0.1792	–160.5595	<0.0001
NY23Ain	6291	46.0066	0.0698	–0.2437	–136.4026	<0.0001
NY23Ain	6291	18.4296	0.0373	–0.1169	–154.7899	<0.0001
NY23Aout	6291	9.9476	–0.0427	0.0876	–102.9699	<0.0001

(continued on next page)

Table C.2. Model Output for Pedal Change Points (continued)

Curve_ID	Radius	beta0	beta1	beta2	ChangePoint	*p*-value
NY23Aout	6291	209.9175	0.4586	–0.4587	–382.2663	0.00455376
NY52Aout	1876	33.6434	0.0844	–0.1067	–39.7973	<0.0001
NY52Aout	1876	84.922	0.0013	–0.0695	–419.8028	<0.0001
NY52Cin	1523	14.2028	–0.0189	0.0776	–200.6215	<0.0001
NY52Cin	1523	67.2325	0.0671	–0.3284	–134.6353	<0.0001
NY52Cout	1523	–71.4701	–0.4301	0.3824	–220.8667	<0.0001
NY52Din	1378	3.1357	–0.0456	0.0462	–179.0635	<0.0001
NY52Dout	1378	50.6786	0.0738	–0.2036	–141.275	0
NY60Ain	1218	48.8336	0.1109	–0.1335	–321.1955	<0.0001
NY60Ain	1218	50.5367	0.1285	–0.1252	–322.8089	0.00316051
NY62Aout	1926	–164.294	–0.5037	0.5336	–387.2163	<0.0001
NY62Aout	1926	18.7174	0.0293	–0.0174	–265.1532	<0.0001
NY63Aout	1357	25.2929	0.026	–0.0513	–243.7218	<0.0001
NY63Aout	1357	15.7151	–0.0024	–0.1413	–59.2882	<0.0001
NY63Aout	1357	28.1578	–0.0243	0.0742	–168.3938	<0.0001
NY63Aout	1357	23.302	–0.0072	–0.0204	–126.3282	<0.0001
NY63Aout	1357	36.184	–0.0019	–0.1122	–53.1712	<0.0001
NY63Aout	1357	21.2713	–0.0068	0.0085	–229.9787	<0.0001
NY65Ain	828	3.3514	–0.0574	0.0572	–278.706	0.18727446
NY65Ain	828	12.8499	–0.0545	0.126	–132.9546	<0.0001
NY65Aout	828	21.7495	–0.0405	0.0651	–206.7026	<0.0001
NY65Aout	828	7.5755	–0.0712	0.0867	–338.4887	<0.0001
NY65Aout	828	24.3846	0.01	–0.1103	–38.2489	<0.0001
NY65Bin	1175	30.1519	–0.0261	0.1202	–57.7677	<0.0001
NY65Bin	1175	38.1532	0.0123	–0.0749	–130.9903	0.00084228
NY65Bin	1175	58.6243	0.1086	–0.1448	–261.9502	<0.0001
NY65Bout	1175	111.8793	0.1479	–0.166	–488.9854	<0.0001
NY65Bout	1175	33.5212	0.0024	–0.3515	–38.8942	0
NY67Aout	117	15.3462	–0.001	0.0593	–118.9137	<0.0001
NY67Aout	117	16.5768	0.0039	0.0211	–159.264	<0.0001
NY67Aout	117	15.5465	–4.00E-04	0.0396	–112.6097	<0.0001
NY67Aout	117	17.3289	0.0086	0.0204	–161.63	<0.0001
NY67Aout	117	20.3922	0	–0.0013	–123.4492	<0.0001
NY67Aout	117	19.0213	–0.0156	1.2858	–30.8386	<0.0001
NY67Aout	117	15.3691	–8.00E-04	0.0839	–87.959	<0.0001

Related SHRP 2 Research

Naturalistic Driving Study: Development of the Roadway Information Database (S04A)

Design of the In-Vehicle Driving Behavior and Crash Risk Study (S05)

Naturalistic Driving Study: Technical Coordination and Quality Control (S06)

Naturalistic Driving Study: Collecting Data on Cell Phone Use (S06)

Naturalistic Driving Study: Field Data Collection (S07)

Analysis of Naturalistic Driving Study Data: Safer Glances, Driver Inattention, and Crash Risk (S08A)

Analysis of Naturalistic Driving Study Data: Offset Left-Turn Lanes (S08B)

Naturalistic Driving Study: Descriptive Comparison of the Study Sample with National Data (S31)

Naturalistic Driving Study: Alcohol Sensor Performance (S31)

Naturalistic Driving Study: Linking the Study Data to the Roadway Information Database (S31)